



Business Intelligence Aplicado a Dados Obstétricos dos Hospitais

ANA FILIPA SOUSA ALMEIDA

Outubro de 2017

Business Intelligence Aplicado a Dados Obstétricos dos Hospitais

Ana Filipa Sousa Almeida

**Dissertação para obtenção do Grau de Mestre em
Engenharia Informática, Área de Especialização em
Sistemas de Informação e Conhecimento**

Orientador: Doutor Constantino Martins

Co-Orientador: Doutor Ricardo Correia, Mestre Tiago Silva Costa

Júri:

Presidente:

Vogais:

Dedicatória

Dedico esta dissertação os meus pais.

Resumo

Um serviço de obstetrícia armazena dados referentes ao percurso das mães e dos recém-nascidos durante todo o período de gestação. Todos estes dados são importantes para que os profissionais de saúde efetuem a sua prática clínica de uma forma mais adequada melhorando assim a prestação dos cuidados de saúde. Existe um conjunto de indicadores importantes para a verificação da qualidade, destacando-se a taxa de cesarianas. Este indicador em Portugal ainda apresenta um valor elevado. O último valor conhecido é referente a 2015 com uma percentagem de 32,9. A execução de uma cesariana tem diversas complicações para a saúde da mãe e/ou para a saúde do recém-nascido, o que leva a que seja muito importante reduzir esta taxa. Devido à sua importância são aplicadas medidas sancionatórias económicas às instituições que apresentem taxas elevadas.

Com vista à redução dessa taxa, é de extrema importância a disponibilização de informação que permita identificar as causas da mesma. Para os profissionais de saúde, é fulcral o acesso aos dados de forma rápida para que possam inferir quais são causas do valor da taxa e atuarem em conformidade. Com este fluxo de conhecimento é possível reduzir a taxa de forma mais eficiente e assim melhorar a qualidade de um serviço de obstetrícia. Atualmente os serviços de obstetrícia não possuem sistemas de informação que disponibilizem o tratamento dos dados de forma a fornecer-lhes a informação necessária para reduzirem a taxa de cesarianas.

Para colmatar esta necessidade e auxiliar os profissionais de um serviço de obstetrícia foi desenvolvido um sistema de *business intelligence* que permite centralizar os dados de várias instituições de saúde e proceder à sua análise. Os indicadores apresentados pela ferramenta foram definidos em conjunto com os profissionais de saúde para que o resultado final fosse de encontro às suas necessidades.

A ferramenta desenvolvida disponibiliza um conjunto de relatórios para que os profissionais possam efetuar as suas análises e acompanhar a evolução destes indicadores. A solução permite ao utilizador a filtragem de informação que é disponibilizada sob a forma de diferentes *dashboards*, dando ao utilizador a possibilidade de geração de relatórios em diferentes formatos. Um dos relatórios disponibilizado tem a comparação da taxa de cesarianas entre hospitais, esta informação permite ao utilizador comparar o serviço de obstetrícia de hospitais diferentes.

Como forma de validação da ferramenta desenvolvida foram elaborados inquéritos de satisfação, onde os utilizadores avaliam as funcionalidades e verificam a sua utilidade no serviço de obstetrícia. Cerca de 90% dos utilizadores inquiridos indicaram que a utilização da solução é intuitiva e simples. Relativamente à aceitação da aplicação, todos os inquiridos mostraram interesse em utilizar aquela aplicação no seu serviço de obstetrícia.

Palavras-chave: Business intelligence, serviço de obstetrícia, indicadores, taxa de cesarianas, relatórios

Abstract

An obstetrics service stores data on the course of mothers and newborns throughout the gestation period. All these data are important for healthcare professionals to perform their clinical practice in a more appropriate way, thus improving the delivery of health care services. There is a set of important indicators for the verification of the quality, being highlighted the cesarean rate. This indicator still presents a high value in Portugal, being the last known value of 32,9 per cent referring to 2015. Performing a cesarean section has several complications both for the health of the mother and of the newborn, which makes it extremely important to reduce this rate. Due to its importance, economic sanctions are applied to institutions with high rates.

Aiming at reducing this rate, providing information is extremely important to allow identifying its causes. For health professionals, to access data quickly is crucial, so that they can infer the causes of the rate value and act accordingly. With this flow of knowledge, it is possible to reduce the rate more efficiently, thus improving the quality of the obstetrical services. Currently, these services do not have information systems that provide the data treatment, to give them the information they need in order to reduce the cesarean rate.

To meet this need and assist the professionals of an obstetrical service, a business intelligence system was developed, which allows the centralization of data coming from several health institutions and their analysis. The indicators presented by the application were defined together with the healthcare professionals, so that the result would meet their needs.

The application developed provides a set of reports to help professionals perform the analyzes and monitor the evolution of these indicators. The solution allows the user to filter the information that is available in the form of different dashboards, giving them the possibility of generating reports in different formats. One of the reports provided has a comparison of the cesarean rate among hospitals, this information allows the user to compare the obstetrical services of different hospitals.

As a way of validating the application developed, satisfaction surveys were elaborated, in which users evaluate the functionalities and verify the usefulness of the application in the obstetrics service. About 90 % of users surveyed indicated that the use of the solution is intuitive and simple. Regarding the acceptance of the application, all respondents showed an interest in using that application in their obstetrics service.

Keywords: Business intelligence, obstetrics service, indicators, cesarean rate, reports.

Agradecimentos

Um especial agradecimento aos meus pais pelo apoio incondicional e por todos os ensinamentos. À minha irmã por toda a motivação e nunca me deixar desistir. Ao meu namorado, Miguel, por todo o apoio, dedicação, paciência e por todos os seus conselhos.

Agradeço aos meus amigos por todo o apoio e motivação. Obrigada por estarem aí, por mostrarem que tudo é possível. Quero destacar o grupo de amigos que fiz durante este mestrado, Bruno Gomes, Nuno Lima, Ricardo Costa e Viviana Batista. Obrigada pela vossa amizade, e pelo companheirismo ao longo destes dois anos.

Agradeço ao meu orientador do ISEP, professor Constantino Martins, por todo o apoio e por todos os seus conselhos para a elaboração desta dissertação.

Quero agradecer ao professor Ricardo Correia, ao Tiago Costa e ao Pedro Farinha pelo apoio e por todo o acompanhamento ao longo desta dissertação.

Agradeço também à VirtualCare pela oportunidade de elaboração desta dissertação.

"Do whatever you will, but first be such as are able to will."

— Friedrich Nietzsche

Conteúdo

Lista de Figuras	xv
Lista de Tabelas	xvii
Lista de Acrónimos	xix
Glossário	xxi
1 Introdução	1
1.1 Enquadramento	1
1.2 Problema	2
1.3 Motivação	3
1.4 Contribuições	3
1.5 Objetivos	4
1.6 Estrutura do documento	4
2 Contexto e Estado da Arte	7
2.1 Descrição do problema	7
2.1.1 Pergunta chave do problema	10
2.2 Análise de Valor	10
2.2.1 New Concept Development Model (NCD)	10
Identificação da oportunidade	11
Análise da oportunidade	11
Geração e enriquecimento de ideias	11
Seleção de ideias	11
Definição do conceito	11
2.2.2 Valor, Valor para o cliente e valor percebido	12
2.2.3 Proposta de valor	13
2.2.4 Modelo de negócio CANVAS	13
2.3 Business Intelligence	15
2.3.1 Processo de <i>Extract, Transform, Load</i> (ETL)	16
2.3.2 <i>Data Warehousing</i>	17
Arquiteturas do <i>Data Warehouse</i> (DW)	17
2.4 Business Intelligence na Saúde	19
2.5 Serviço de Obstetrícia	23
2.6 Conclusão	24
3 Ferramentas de Business Intelligence	25
3.1 Análise das ferramentas	25
3.1.1 Pentaho	27
3.1.2 TIBCO software	28
3.1.3 Microsoft	28

3.1.4	Birt	29
3.1.5	SpagoBI	29
3.2	Comparação das ferramentas	30
3.2.1	Grandezas a Avaliar	32
3.3	Conclusão	35
4	Design da solução proposta	37
4.1	Introdução	37
4.2	Requisitos	38
4.2.1	Requisitos não funcionais	38
4.2.2	Requisitos funcionais	39
4.3	Arquiteturas propostas	40
4.4	Arquitetura implementada	43
4.4.1	Modelação do DW	45
4.4.2	Implementação do DW	53
	Carregamento das dimensões	53
	Carregamento da tabela de factos	55
4.4.3	Configuração da visualização dos dados	56
4.5	Conclusão	60
5	Análise dos resultados	61
5.1	Indicadores	61
5.1.1	Número de cesarianas anual	61
5.1.2	Número de partos realizados	65
5.1.3	Motivos de uma cesariana	66
5.1.4	Complicações das cesarianas	67
5.1.5	Recém nascidos	69
5.1.6	Caracterização da mãe em cesarianas	74
5.1.7	Tipo de episódios	75
5.1.8	Comparação entre instituições	75
5.2	Conclusão	76
6	Avaliação da Solução	79
6.1	Metodologia de avaliação	79
6.2	Resultados do inquérito disponibilizado aos médicos	80
6.3	Resultados do inquérito disponibilizado aos colaboradores da VirtualCare	81
6.4	Conclusão	82
7	Conclusão	83
7.1	Desenvolvimentos Futuros	84
A	Método AHP para escolha da ferramenta	91
B	Inquérito apresentado aos médicos	95
C	Inquérito apresentado aos colaboradores da VirtualCare	97

Lista de Figuras

Figura 2.1	Modelo de negócio, CANVAS.	14
Figura 2.2	Arquitetura de um sistema de <i>Business Intelligence</i> (BI) (Chaudhuri, Dayal e Narasayya, 2011).	15
Figura 2.3	Arquitetura de Inmon.	18
Figura 2.4	Arquitetura de Kimball.	18
Figura 2.5	Arquitetura conceptual de um sistema de BI para o Intermountain HealthCare (Evans, Lloyd e Pierce, 2012).	21
Figura 2.6	Arquitetura conceptual do sistema de BI aplicado ao grupo de diagnósticos homogéneos (Barrento et al., 2010).	22
Figura 3.1	Quadrante Mágico (Parenteau et al., 2016).	27
Figura 3.2	Esquematisação do problema, dos critérios e das alternativas.	33
Figura 4.1	Diagrama de casos de uso.	39
Figura 4.2	Diagrama de fluxo de dados para a arquitetura proposta.	41
Figura 4.3	Diagrama de componentes da arquitetura proposta.	42
Figura 4.4	Vista concetual de implementação.	42
Figura 4.5	Diagrama de fluxo de dados para a alternativa de arquitetura proposta.	43
Figura 4.6	Diagrama de componentes para a arquitetura alternativa.	43
Figura 4.7	Relação entre a dimensão da data e a tabela de factos.	46
Figura 4.8	Relação entre a dimensão do local e a tabela de factos.	47
Figura 4.9	Relação entre a dimensão dos tipos de documento e a tabela de factos.	47
Figura 4.10	Relação entre a dimensão dos tipos de parto e a tabela de factos.	48
Figura 4.11	Relação entre a dimensão dos motivos e a tabela de factos.	48
Figura 4.12	Relação entre a dimensão dos tipos de episódio e a tabela de factos.	49
Figura 4.13	Relação entre a dimensão das complicações e a tabela de factos.	49
Figura 4.14	Relação entre a dimensão dos tipos de gravidez e a tabela de factos.	49
Figura 4.15	Relação entre a dimensão da apresentação e a tabela de factos.	50
Figura 4.16	Relação entre a dimensão dos doentes e a tabela de factos.	50
Figura 4.17	Modelo do <i>Data Mart</i> (DM) implementado para o serviço de obstetrícia	52
Figura 4.18	Escolha da fonte que possui os dados, base de dados oracle.	57
Figura 4.19	Autenticação da base de dados.	57
Figura 4.20	Tabelas disponíveis para visualização no Power BI.	57
Figura 4.21	Relações entre as tabelas representadas no Power BI.	58
Figura 4.22	Relatório principal de taxa de cesarianas partilhado no <i>sharepoint</i>	60
Figura 5.1	Relatório com taxas anuais fornecido ao utilizador.	62
Figura 5.2	Percentagem de cesarianas para o primeiro semestre do ano corrente.	63
Figura 5.3	Taxa de cesarianas por trimestre.	63
Figura 5.4	Taxa de cesarianas por mês.	64
Figura 5.5	Relatório com o número de partos caracterizados pelo tipo de parto.	65
Figura 5.6	Relatório com os motivos das cesarianas.	67

Figura 5.7	Relatório com as complicações das cesarianas.	68
Figura 5.8	Relatório com o número de recém nascidos.	69
Figura 5.9	Relatório com o número de recém nascidos do ano 2017, do sexo masculino.	70
Figura 5.10	Relatório com o número de recém nascidos distribuídos por ano e por mês.	71
Figura 5.11	Filtragem da distribuição de recém nascidos por mês para o ano de 2016.	71
Figura 5.12	Filtragem da distribuição de recém nascidos para o mês de janeiro, por ano.	72
Figura 5.13	Relatório com o número de cesarianas por apresentação do bebê. . .	73
Figura 5.14	Relatório do número de cesarianas caracterizadas pelo peso e idade da mãe.	74
Figura 5.15	Relatório com o número de episódios diferenciados pelo seu tipo. . .	75
Figura 5.16	Relatório com a taxa de cesarianas dos diferentes hospitais.	76
Figura 5.17	Apresentação das opções de execução do relatório fornecidas pelo Power BI.	77

Lista de Tabelas

Tabela 2.1	Características das arquiteturas de Inmon e Kimball.	19
Tabela 2.2	Indicadores do serviço de obstetrícia.	23
Tabela 3.1	Verificação das características das ferramentas encontradas.	31
Tabela 3.2	Comparação dos pesos de importância.	33
Tabela 6.1	Escala utilizada nos inquéritos.	79
Tabela 6.2	Frequência relativa de cada resposta do inquérito disponibilizado aos médicos.	80
Tabela 6.3	Frequência relativa de cada resposta do inquérito aos colaboradores da VirtualCare.	81
Tabela A.1	Avaliação das ferramentas para o critério escalabilidade.	91
Tabela A.2	Avaliação das ferramentas para o critério configuração.	92
Tabela A.3	Avaliação das ferramentas para o critério performance.	92
Tabela A.4	Avaliação das ferramentas para o critério integração.	92
Tabela A.5	Apresentação das prioridades relativas e da prioridade composta. . .	93

Lista de Acrónimos

ARS	Administração Regional de Saúde.
ARS Norte	Administração Regional de Saúde do Norte.
BD	Base de Dados.
BI	<i>Business Intelligence.</i>
DGS	Direção-Geral de Saúde.
DM	<i>Data Mart.</i>
DW	<i>Data Warehouse.</i>
ETL	<i>Extract, Transform, Load.</i>
OLAP	<i>Online Analytical Processing.</i>
SA	<i>Staging Area.</i>
SIM@SNS	Sistema Informação Monitorização do Serviço Nacional de Saúde.
URL	<i>Uniform Resource Locator.</i>

Glossário

<i>Business Intelligence</i>	Sistema de informação de suporte na tomada de decisão de uma área de negócio.
<i>Dashboards</i>	Apresentação dinâmica da informação.
<i>Data Mart</i>	Subconjunto de dados, modelado de forma multidimensional orientado a departamentos específicos das organizações..
<i>Data Warehouse</i>	Sistema de armazenamento de dados de uma área de negócio. Base de dados utilizada para os processos de análise dos dados.
<i>MapReduce</i>	Divisão dos dados, e distribuição pelos vários processos. Processo utilizado na agregação de dados e pesquisas.
<i>Online Analytical Processing</i>	Processos que permitem manipular e analisar a informação de múltiplas perspectivas.
<i>Staging Area</i>	Área de armazenamento onde ocorrem as transformações dos dados. Camada antes do <i>Data Warehouse</i> .
Histerectomia	Operação cirúrgica para remoção do útero.

Capítulo 1

Introdução

No presente capítulo é elaborado um enquadramento do serviço de obstetrícia do sistema de saúde português onde são apresentadas as principais necessidades bem como o problema que enfrentam. Ao longo do capítulo é apresentada a motivação para a elaboração desta dissertação, assim como, a forma como esta vai contribuir na resolução do problema encontrado. São também apresentados os objetivos para esta dissertação.

1.1 Enquadramento

A utilização dos sistemas de informação na área da saúde levou a que mais facilmente se possa armazenar os dados, no entanto, as instituições de saúde têm de garantir, cada vez mais, uma melhor prestação de cuidados de saúde sendo para isso necessário que toda a informação dos utentes (clínica e pessoal) esteja guardada e facilmente acessível (Dale, 2014).

Num serviço de obstetrícia existe a necessidade de armazenar dados de forma a que se possa acompanhar o percurso da mãe e do seu bebé (Peng et al., 1992). Os dados mais relevantes que devem ser armazenados são: o histórico de outras gestações (tipo de parto, principais complicações); exames realizados durante a gestação; histórico familiar (doenças genéticas). Estes dados apresentam uma mais valia para o profissional de saúde permitindo-lhe uma análise mais rápida da informação dispondo assim de mais tempo para avaliar e tomar uma decisão clínica.

As constantes mudanças socio-culturais aliadas ao crescente progresso da tecnologia levam a que a escolha pelo tipo de parto (vaginal ou cesariana) a realizar seja encarada de forma diferente. A mãe pode já ter uma ideia formada sobre como quer o parto e pode ser difícil fazê-la mudar de opinião. Caso se decida por uma cesariana é necessário informar a mãe os riscos associados não só para ela, como também para o bebé (Mylonas e Frieze, 2015a).

A Administração Regional de Saúde do Norte (ARS Norte), instituição que efetua a administração das políticas da saúde da área Norte adquiriu para todos os seus hospitais a aplicação informática ObsCare (*VC ObsCare – VirtualCare – Systems for Life*). Esta aplicação é utilizada em todos os hospitais do Norte para efetuar o registo clínico eletrónico das utentes no serviço de obstetrícia. Com esta ferramenta é possível obter toda a informação da utente relativa à sua gestação, complementando essa informação com os vários exames que a paciente necessitou de fazer durante a gravidez. A taxa de cesarianas é um dos indicadores que é necessário obter da informação guardada, assim como os motivos pelos quais foi efetuada, para que assim se possa verificar se a cirurgia era de facto necessária.

A escolha para efetuar uma cesariana pode ser devido a vários fatores, por exemplo, patologias maternas, anomalias fetais e patologias próprias da gravidez (George, 2015). No entanto, existem cesarianas que não são efetuadas por motivos clínicos, nem pela vontade da mãe, mas sim porque permite ao médico programar a sua agenda por ser mais fácil de prever o tempo de duração de uma intervenção cirúrgica. Outro dos motivos é porque se sente mais confortável a realizar esse procedimento (Hannah Da Silva Copelli et al., 2015; Faisal-Cury e Menezes, 2006).

Em alguns casos os médicos indicam argumentos clínicos como motivo para efetuar a cesariana, situações em que não haveria necessidade de uma cesariana. O profissional aproveita-se do desconhecimento da mãe ou mesmo da sensibilidade do momento apresentando motivos que não são os corretos para poder convencer a mãe que é a melhor opção (Hannah Da Silva Copelli et al., 2015; Faisal-Cury e Menezes, 2006). Apesar da mãe saber todos os prós e contras da escolha de uma cesariana acaba por seguir o "conselho" do médico, uma vez que é ele o responsável por indicar o tipo de parto.

Os motivos pelos quais é efetuada uma cesariana são importantes para se poder efetuar auditorias clínicas e para se ter o conhecimento claro da decisão tomada pelo médico (George, 2015), como tal, os hospitais têm de efetuar a monitorização dos cuidados prestados no setor de obstetrícia, acompanhando o número de cesarianas e quais os motivos para efetuarem essa intervenção.

Em Portugal a Direção-Geral de Saúde (DGS) definiu que as cesarianas deveriam ser classificadas segundo motivos principais (George, 2015) (por exemplo, cirurgia uterina prévia; trabalho de parto estacionário), isso leva a que os médicos tenham de enquadrar a cesariana num desses motivos.

Uma cesariana apresenta, para as instituições de saúde, um custo superior, em comparação com um parto vaginal, não só por se tratar de uma cirurgia mas também porque pode levar a complicações clínicas tanto para a mãe como para o bebé, o que leva a que estes necessitem de um maior número de cuidados de saúde.

A taxa de cesarianas é considerada internacionalmente como o principal indicador de cuidados obstétricos, em Portugal o último valor calculado da taxa é de 32,9 % para 2015. Como tal existem incentivos financeiros para reduzir o número de cesarianas em hospitais públicos e foi criada uma comissão com vista a atingir este objetivo.

Assim sendo, é necessário informar os profissionais de saúde para que estes tomem melhores decisões clínicas permitindo assim a redução da taxa.

1.2 Problema

Em Portugal, segundo a Norma nº 001/2015 de 19/01/2015 da Direção-Geral de Saúde (George, 2015), os hospitais têm de possuir registos clínicos onde é indicado o motivo da cesariana, quando foi efetuada, qual o médico que optou pela cesariana, entre outros dados. Sendo obrigatório enviar anualmente o número de cesarianas efetuadas em cada hospital, respeitando determinadas condições presentes na norma, indicada anteriormente.

Os hospitais, caso necessário, têm de disponibilizar os dados para auditorias que possam ser efetuadas, internas e externas, como tal precisam de ter uma estrutura onde armazenam os dados e facilmente obtenham essa informação. Atualmente os hospitais pertencentes à

ARS Norte possuem uma aplicação, que efetua a recolha de todos estes dados, no entanto, esta aplicação não apresenta toda a informação que os profissionais de saúde necessitam.

A aplicação não permite o acompanhamento da taxa de cesarianas nem de outros indicadores de forma automática, não fornecendo a informação necessária e atempada para auxiliar os profissionais de saúde no seu processo de tomada de decisão.

1.3 Motivação

O tema apresentado identifica um problema na área da saúde nacional. A escolha deste tema prende-se com o gosto prévio pela área da saúde, assim como, a vontade de contribuir para a melhoria de um sistema indispensável a qualquer cidadão, o sistema de saúde.

O tema surgiu de uma necessidade dos profissionais de saúde, problema esse, que solucionado pode ajudar o quotidiano destes profissionais. Apresentar uma solução que possa ajudar na resolução desse problema torna-se um grande desafio e é bastante estimulante, uma vez que o que vai ser desenvolvido pode facilitar o quotidiano dos médicos, mas mais que isso, pode levar a que sejam prestados melhores cuidados de saúde às utentes.

Na área da saúde existe um grande volume de dados movimentado nos vários setores, no entanto, nem sempre é efetuada a correta exploração dos mesmos. A oportunidade de poder extrair o conhecimento necessário para os médicos também ajudou na escolha deste tema para efetuar esta dissertação.

As ferramentas de *Business Intelligence* (BI) têm vindo a apostar na área da saúde, e uma mais valia pessoal é a possibilidade de analisar diferentes ferramentas de análise e tratamento de dados do mercado. A possibilidade de elaborar a arquitetura para estes sistemas assim como a sua implementação são um fator aliciante para a realização desta dissertação.

1.4 Contribuições

A presente dissertação propõe um sistema de apoio aos profissionais de saúde, baseado num sistema de *Business Intelligence*. As instituições de saúde geram um grande volume de dados, podendo ser divididos em três categorias: financeiros (dados relativos aos custos ou receitas das instituições); administrativos (dados demográficos dos utentes, bem como, dados do funcionamento da instituição); clínicos (dados relativos ao estado de saúde e doença do utente) (Mettler e Vimarlund, 2007; Ashrafi, Kelleher e Kuilboer, 2014). A implementação deste sistema vai permitir realizar a gestão destes, que posteriormente serão utilizados para otimização do serviço de obstetrícia. As principais contribuições para o serviço são a possibilidade de suportar o desempenho clínico, melhorar processos, redução de custos.

Com estes sistemas a informação estará disponível sempre que necessário para os utilizadores poderem utilizar, o que pode melhorar a tomada de decisão, podendo ser tomada mais eficiência. Com o acesso facilitado aos dados do paciente podem ser gerados vários relatórios com base em diferentes variáveis, por exemplo, idade da mãe, tipo de parto.

As instituições podem efetuar análises descritivas com base em factos passados e deteção de padrões, levando a que sejam tomadas as decisões atuais ou futuras baseadas em mais informação (Raghupathi e Raghupathi, 2013).

O sistema proposto irá processar os dados obtidos pelo setor de obstetrícia, garantindo a qualidade e integridade dos dados, para que posteriormente sejam explorados pelo modelo proposto.

Inicialmente é efetuado um estudo e é escolhida a ferramenta que melhor se adequa ao problema encontrado, essa ferramenta será configurada de forma a adaptar-se aos vários clientes. Para que os médicos tenham facilmente acesso à informação, esta será disponibilizada sob a forma de *Dashboards* e relatórios tornando a sua análise o mais intuitiva possível. O modelo desenvolvido irá apresentar um conjunto de indicadores que irão permitir caracterizar o serviço prestado na prática clínica de obstetrícia, bem como, comparações desse serviço nos vários hospitais. O profissional de saúde vai assim tomar a decisão com mais segurança uma vez que terá um conjunto de factos, fornecidos pela ferramenta, que o auxiliam na tomada de decisão.

1.5 Objetivos

O objetivo principal desta dissertação é o desenvolvimento de um sistema de informação de BI que forneça a informação necessária aos profissionais de saúde no seu processo de tomada de decisão com o intuito de melhorar a qualidade dos serviços de obstetrícia.

O sistema proposto terá de ser capaz de explorar dados obstétricos e compará-los (se for o caso) entre as diferentes instituições de saúde através de objetivos definidos pela ARS Norte.

A ferramenta é para utilização dos profissionais de saúde e como tal, tem de ser capaz de gerar *dashboards* para que o utilizador possa verificar os dados atualizados e comparar a taxa de cesarianas entre os vários hospitais, deve também permitir ao utilizador efetuar filtragens tornando a visualização dos dados interativa. Como os dados já se vão encontrar armazenados, para além da taxa de cesarianas poderão ser obtidos indicadores de qualidade, caso o utilizador assim o pretenda. Destes indicadores pode destacar-se, por exemplo, taxa de partos vaginais após cesariana, percentagem do primeiro parto por cesariana (Direção-Geral da Saúde, 2016).

O sistema deverá ainda permitir a execução de relatórios com a informação que o utilizador pretende visualizar.

1.6 Estrutura do documento

O presente documento encontra-se estruturado em sete capítulos. No primeiro capítulo é efetuado um enquadramento do projeto, o problema encontrado, as contribuições para a resolução desse problema e os objetivos.

No segundo capítulo é apresentada a descrição detalhada do problema, é efetuada uma análise de valor do projeto que se vai desenvolver. Neste capítulo são também apresentados o conceito de BI, assim como a estrutura destes sistemas na área da saúde.

No terceiro capítulo é efetuado um estudo das ferramentas de análise existentes no mercado, procedendo-se à avaliação, segundo o método de AHP, dessas ferramentas encontradas.

No quarto capítulo são apresentados os requisitos funcionais e não funcionais sendo ainda propostas duas arquiteturas (principal e alternativa) para o desenvolvimento da solução. Neste capítulo é também descrita a implementação da arquitetura escolhida.

No quinto capítulo são apresentados os resultados obtidos, os diferentes indicadores e é efetuado o mapeamento desses resultados com os requisitos propostos.

No sexto capítulo é apresentada a avaliação da solução desenvolvida, são apresentados os inquéritos efetuados aos utilizadores, assim como os seus resultados.

Por último, o sétimo capítulo apresenta uma conclusão do documento .

Capítulo 2

Contexto e Estado da Arte

No presente capítulo é elaborada uma descrição detalhada do problema, o qual esta dissertação pretende solucionar. São apresentadas as várias etapas do modelo de conceção e desenvolvimento de Koen enquadradas no desenvolvimento da solução proposta. É também apresentado o valor da solução para os seus utilizadores, bem como uma visão global de como se pretende criar e entregar esse valor para o cliente (modelo de negócio CANVAS).

Neste capítulo é também apresentado o estado da arte de forma a esclarecer o que são sistemas de BI, e como é que estes sistemas podem ser uma mais valia para a área da saúde.

Por último é apresentado o serviço de obstetrícia, as suas lacunas e alguns dos indicadores pertinentes para o serviço.

2.1 Descrição do problema

Com os dados recolhidos atualmente no serviço de obstetrícia, torna-se importante o acompanhamento constante por parte dos médicos, no entanto, a solução que atualmente possuem não permite esse acompanhamento, que é importante não só para os médicos dos respetivos hospitais, diretores de serviço, mas também para a ARS Norte. Por exemplo, é relevante para o diretor de serviço saber a percentagem de cesarianas no seu hospital de forma automática, uma vez que esse valor pode comprometer o financiamento do seu serviço.

A ARS Norte agrega os dados dos hospitais do Norte entre os quais dados de obstetrícia que podem ser importantes para a análise dos cuidados de saúde desses hospitais. Atualmente é recolhida uma grande quantidade de dados neste serviço, no entanto, estes não estão a ser utilizados nem analisados por não existirem os recursos suficientes.

Face a esta lacuna, encontrada nos vários serviços de obstetrícia do Serviço Nacional de Saúde (SNS), é proposto um sistema de BI que permita explorar os dados obstétricos dos hospitais do Norte, como caso de estudo será aplicado no Centro Hospitalar de São João. Esta solução permite obter indicadores (taxa de cesarianas) com base na informação fornecida que são importantes quer para os médicos dos hospitais quer para a ARS Norte, tendo esta ultima instituição interesse na comparação de dados inter-hospitalares. Desta forma, é permitido aos profissionais de saúde uma consulta mais rápida da informação que necessitam, disponibilizando assim de mais tempo para dedicarem à tomada de decisões clínicas.

A ferramenta irá aceder a dados do serviço de obstetrícia recolhidos pelo ObsCare, efetuar o seu tratamento e analisar os dados de forma a apresentar aos vários clientes as diferentes estatísticas. A informação será apresentada, pela mesma ferramenta, aos médicos sob a forma de relatórios.

Efetuar uma cesariana tem vantagens e desvantagens, no entanto, é de ressaltar as várias complicações que pode trazer quer para a mãe quer para o bebé. Apesar da evolução das tecnologias e da informação disponível sobre o tema, ainda existem cesarianas que são efetuadas desnecessariamente, podendo ser evitadas. A taxa de cesarianas em Portugal, em 2015, foi de 32,9% (INE, DGS MS e PORDATA, 2016) num total de 83 957 partos efetuados nos estabelecimentos de saúde (INE, DGS/MS e PORDATA, 2017), valores provisórios.

O valor da taxa de cesarianas afeta o financiamento do setor de obstetrícia, isto é, o valor de financiamento é calculado consoante esta percentagem. Se a taxa de cesarianas se encontrar abaixo dos 25% o hospital recebe a totalidade do valor dessa cirurgia. Se a taxa estiver entre os 25% e os 27,5% o hospital recebe 75% do financiamento. Se a taxa se encontrar entre os 27,5% e os 30% é pago apenas metade do valor da cesariana ao hospital. Se a taxa estiver entre 30,0% e 32,5% apenas é dado 0,25% do financiamento que estava previsto e se for superior aos 32,5% o hospital não recebe nenhum financiamento (Campos et al., 2010).

Uma cesariana pode levar a um conjunto de complicações: hemorragias; complicações anestésicas; infeções; trombo-embolismos; entre outras complicações. A escolha por uma cesariana tem de ser efetuada de forma ponderada, se a cesariana não for necessária pode originar-se riscos e complicações desnecessários.

Para os hospitais, a diminuição da taxa de cesariana é extremamente importante, não apenas por questões de saúde das mulheres, mas também pelos valores económicos que estão subjacentes a essa taxa. Nos casos da taxa ultrapassar os 32,5% o hospital pode mesmo não receber qualquer financiamento.

Atualmente não é possível obter de forma automática uma análise detalhada dos dados guardados, deste modo, é importante disponibilizar uma ferramenta que possa efetuar todas estas análises podendo ajudar os profissionais de saúde na monitorização destas taxas.

A ferramenta desenvolvida será configurável, e como tal poderá ter como potencial cliente todos os hospitais do Norte que tenham o ObsCare instalado, uma vez que é esta aplicação que faz toda a recolha dos dados sendo depois analisados e apresentados pela solução proposta. Os hospitais que já possuem o ObsCare instalado são:

- Centro Hospitalar de Entre o Douro e Vouga, E.P.E;
- Centro Hospitalar De São João, E.P.E;
- Unidade Local de Saúde de Matosinhos;
- Hospital da Senhora da Oliveira - Guimarães;
- Centro Hospitalar de Trás-os-Montes e Alto Douro, EPE;
- Centro Hospitalar Póvoa de Varzim - Unidade de Vila do Conde;
- Centro Hospitalar Tâmega e Sousa;
- Unidade Local de Saúde do Nordeste;

- Centro Hospitalar Gaia/Espinho;
- Unidade Local de Saúde do Alto Minho;
- Hospital Geral de Santo António;
- Centro Hospitalar Médio Ave.

Outro potencial cliente poderá ser a ARS Norte, uma vez que esta organização efetua a agregação dos dados de todas as instituições de saúde do Norte, espera-se que com a solução proposta, possa obter relatórios comparativos das várias instituições. A solução proposta também vai facilitar na obtenção de dados estatísticos mais confiáveis, uma vez que os dados estarão centralizados e acessíveis à solução, o que vai permitir à ARS Norte analisar os cuidados prestados no setor de obstetrícia em cada hospital. Irá permitir monitorizar a taxa de cesarianas nos vários hospitais, efetuar comparações interinstitucionais, entre outros indicadores.

Com a utilização deste produto os clientes poderão analisar os dados, por exemplo, os médicos poderão filtrar a informação com base em algumas variáveis (variáveis previamente definidas por eles), terão também a possibilidade de monitorizar as taxas de cesariana, do seu hospital, sob a forma de *dashboards*, verificando se estão a ser tomadas todas as medidas necessárias para que essa taxa seja o mais baixa possível. A ARS Norte poderá também efetuar essa monitorização mas de todos os hospitais do Norte, verificando assim quem apresenta as maiores taxas podendo tomar alguma medida para minimizar estes valores, ou mesmo notificá-los para que passem a ter um maior cuidado.

Para o desenvolvimento do sistema proposto não será desenvolvida uma nova tecnologia, será utilizada uma ferramenta existente de BI para analisar os dados. Esta ferramenta será adaptada e configurada para cada cliente, para melhor responder às necessidades de cada um.

Existem várias partes interessadas para o desenvolvimento desta solução que são:

- VirtualCare, spin-off da universidade do Porto que comercializa o ObsCare;
- Hospitais pertencentes à ARS Norte (vão poder usufruir de uma monitorização constante dos valores recolhidos pelo ObsCare);
- ARS Norte, poderá usufruir da análise dos indicadores de obstetrícia, fornecidos pela solução.

Todas estas entidades têm interesse no desenvolvimento, uma vez que lhes serão fornecidos dados automáticos para o suporte à prática clínica.

Os dados utilizados pelo sistema são recolhidos pelo ObsCare, que efetua o seu armazenamento numa base de dados Oracle, onde o protótipo terá de aceder aos dados. Esta aplicação recolhe os dados referentes ao período de gestação das mulheres, por exemplo, guarda a data em que estas fazem as ecografias, análises ou outros exames que sejam necessários. A aplicação permite efetuar um registo diário com todos os dados relevantes para o acompanhamento da gestação da mulher. Posto isto, não será necessário utilizar outra fonte de informação para obtenção dos dados utilizados no projeto, uma vez que esta aplicação tem todos os dados necessários armazenados. No entanto, a solução desenvolvida só poderá ser utilizada nos hospitais que tenham o ObsCare instalado e na ARS Norte.

A solução desenvolvida vai comunicar com outros sistemas, nomeadamente a base de dados que efetua o armazenamento dos dados recolhidos pelo ObsCare. A arquitetura da solução deverá ter em conta estas comunicações, uma vez que estamos a lidar com dados críticos, deverá utilizar-se comunicações seguras para garantir toda a integridade e confidencialidade dos dados.

2.1.1 Pergunta chave do problema

Com o problema encontrado coloca-se uma pergunta chave à qual o sistema proposto deverá responder, que é: Como é possível reduzir a taxa de cesarianas com o fornecimento automático, de informação referente aos vários indicadores e à taxa de cesarianas para os profissionais de saúde?

2.2 Análise de Valor

Para validar a viabilidade da solução foi elaborada uma análise de valor, que é apresentada no presente capítulo. Esta análise permite verificar os benefícios para o utilizador.

2.2.1 New Concept Development Model (NCD)

Para se desenvolver um produto final é necessário seguir algumas fases para todo este processo de inovação e de desenvolvimento do produto. O processo de inovação tem contemplada três fases: Fuzzy Front-End (FFE), New Product Development (NPD), Comercialização. A fase de FFE é considerada como uma grande oportunidade para melhorar o processo de inovação, uma metodologia utilizada no FFE é o NCD, apresentado por Peter Koen (Methods et al., 1996). Este modelo fornece uma linguagem comum para as atividades de Front-End, o que vai permitir uma partilha e uma criação de novo conhecimento. Este modelo é dividido em três partes: o motor que vai potenciar os elementos (visão, estratégia que impulsiona a inovação, cultura corporativa para estimular a criatividade, a criação de conhecimento); os cinco elementos Front-End; e os fatores externos.

Os cinco elementos Front-End são: Identificação da oportunidade; Análise da oportunidade; Geração e enriquecimento de ideias; Seleção de ideias; Definição do conceito. Um projeto pode começar por duas formas: geração de ideias, fase onde são apresentadas ideias e é a partir daí que é desenvolvido um produto; ou pela identificação da oportunidade no mercado, onde é verificada uma lacuna e é desenvolvido um produto que colmata essa lacuna.

O tema apresentado nesta dissertação surgiu de uma lacuna no setor de obstetrícia, foi identificada uma oportunidade, verificando-se que existiam dados suficientes para se fazer análises e apoiar os profissionais de saúde que não estavam a ser devidamente usados. As secções seguintes apresentam os vários elementos do modelo NCD e quais as técnicas utilizadas em cada um.

Identificação da oportunidade

A oportunidade surgiu de um problema existente no setor de obstetrícia, onde se verificou que existia uma lacuna no tempo de resposta e na quantidade de informação fornecida aos profissionais de saúde. Usufruidos da evolução das tecnologias e da utilização dessa tecnologia para efetuar a recolha dos dados, analisou-se a oportunidade de utilizar esses dados guardados para melhorar o serviço prestado pelos profissionais. Para comprovar que de facto se tratava de uma oportunidade, e de uma boa oportunidade, verificou-se que era um problema que ocorria nos vários hospitais do serviço nacional de saúde.

Análise da oportunidade

Existem 18 hospitais pertencentes à ARS Norte (Administração Regional de Saúde do Norte, 2017), destes hospitais o ObsCare encontra-se instalada em 12, tal como referido na secção 2.1. Nos hospitais que já possuem o ObsCare para a recolha dos dados, constatou-se que se poderia implementar um sistema de informação para dar resposta ao problema encontrado, uma vez que a recolha dos dados já estava a ser efetuada. Uma perspetiva da empresa detentora do ObsCare é alargar este número de hospitais passando para as outras Administrações Regionais de saúde (Centro, Lisboa e Vale do Tejo, Alentejo e Algarve), esta expansão leva a que esta oportunidade se torne ainda mais atraente.

Para além de toda esta análise foi necessário estabelecer o contacto direto com os profissionais de saúde verificando-se quais as suas tendências e o que eles pretendem.

Geração e enriquecimento de ideias

Em conjunto com os vários intervenientes do projeto efetuou-se um *brainstorming* como forma de apresentar várias ideias para solucionar o problema encontrado. Periodicamente foram realizadas reuniões para incentivar esta partilha de ideias para se perceber qual o sistema que melhor se adequa e quais as ferramentas que se podem enquadrar nesses sistemas.

Depois de identificado o tipo de sistema a utilizar, foram surgindo várias ideias no que diz respeito a como esse sistema deverá aceder aos dados guardados pelo ObsCare e como os dados poderiam ser apresentados aos profissionais de saúde.

Seleção de ideias

Através das várias ideias que surgiram, verificou-se qual seria a mais atrativa, constatando-se que a elaboração de um sistema de BI era a mais viável e que melhor se enquadrava no problema encontrado. O setor de obstetrícia apresenta um grande volume de dados para analisar, como tal podem, e devem, ser obtidos vários indicadores que informem como o serviço de obstetrícia é prestado.

Definição do conceito

A cesariana é uma cirurgia tendo riscos associados e complicações não só para a mãe como para o bebé. Para além das várias complicações de saúde que podem surgir, como se

trata de uma cirurgia são necessários mais recursos (financeiros, recursos físicos, recursos humanos) em comparação com o nascimento de um bebé por parto vaginal. O hospital tem de despende de mais recursos económicos tornando-se este método mais caro. Em Portugal a taxa de cesarianas tem vindo a diminuir, no entanto, ainda continua com valores elevados atingindo 32,9% em 2015 (valor ainda provisório).

Uma das medidas tomadas para reduzir a taxa de cesarianas foi relativamente ao financiamento desta cirurgia, passando-se a condicionar este financiamento com base na taxa de cesarianas. É necessário tomar todas as medidas para reduzir estes valores, e para isso é apresentado um sistema que ajude nesse processo.

O sistema que deverá ser elaborado é um sistema de informação de BI, tal como referido anteriormente, que terá de aceder aos dados que se encontram numa base de dados Oracle. A solução terá de combinar a recolha de dados operacionais e posterior armazenamento num repositório adequado onde deverá ser efetuada toda a transformação dos dados. A gestão destes dados terá de ser efetuada com recurso a ferramentas de análise e exploração dos dados para que posteriormente a informação seja disponibilizada aos utilizadores, fornecendo-lhes a informação necessária para tornarem a sua tomada de decisão mais eficiente e assertiva.

Com as ferramentas de exploração e análise de dados irão obter-se indicadores que vão permitir analisar a qualidade dos cuidados obstétricos, e também comparação interinstitucional. O sistema vai fornecer relatórios, podendo o utilizador usar algumas variáveis para filtrar esses dados. A apresentação da informação ao utilizador vai ser efetuada da forma mais perceptível possível, o que vai permitir ao profissional de saúde analisar rapidamente a informação o que lhe dará um suporte maior na sua tomada de decisão.

2.2.2 Valor, Valor para o cliente e valor percebido

O valor é a chave do marketing e alguns autores definem-no do seguinte modo: “Value has been defined in different theoretical contexts as need, desire, interest, standard /criteria, beliefs, attitudes, and preferences.” (NICOLA, Ferreira e Pinto Ferreira, 2012)

O conceito valor é subjetivo, dependendo de vários fatores, não estando apenas associado a valores económicos, mas também aos benefícios que o cliente consegue obter ou o seu interesse por um determinado produto ou serviço. O valor que se consegue obter vai depender dos benefícios e dos sacrifícios para obter determinado produto ou serviço.

No entanto, o valor para o cliente vai depender das suas necessidades e expectativas. Alguns autores definem valor para o cliente da seguinte forma:

“Value for the customer (VC) is any demand-side, personal perception of advantage arising out of a customer’s association with an organisation’s offering, and can occur as reduction in sacrifice; presence of benefit (perceived as either attributes or outcomes); the resultant of any weighed combination of sacrifice and benefit; or an aggregation, over time, of any or all these.” (Woodall, 2003)

Nem todos os clientes percebem o valor de um determinado produto da mesma forma, e como tal, o valor percebido pelo cliente pode ser uma avaliação da sua utilidade baseando-se nos benefícios e sacrifícios desse determinado produto. Alguns autores defendem que a suscetibilidade para o preço depende de cliente para cliente afetando assim a percepção de

valor desse produto para os vários clientes (Lindgreen e Wynstra, 2005), isto pode assim alterar, para as organizações, o modo como entregam valor para o cliente.

O objetivo inicial para efetuar a análise de valor é verificar como é possível aumentar o valor do produto a um baixo custo, de forma a não comprometer a sua qualidade.

Nesta dissertação o produto que vai ser entregue vai ter vários benefícios mas também alguns pontos negativos. A ferramenta criada vai ser utilizada por profissionais de saúde fornecendo-lhes os factos necessários para suportarem a decisão que vão tomar, a solução proposta vai assim diminuir o tempo que eles precisam para analisar dados anteriores, uma vez que a ferramenta já o vai efetuar.

A solução proposta é uma ferramenta informática, e como qualquer outra, vai necessitar de uma curva de aprendizagem, por mais pequena que seja. Para além da formação inicial vão existir outros custos associados, como por exemplo os custos associados ao alojamento da aplicação.

2.2.3 Proposta de valor

A solução apresentada será uma ferramenta de análise de dados voltada para os profissionais de saúde quer dos hospitais quer da ARS Norte. Os profissionais vão ter acesso a um sistema de apoio à decisão que lhes fornece um conjunto de indicadores, fundamentais para a sua prática clínica, por exemplo, taxa de cesariana.

Com a solução proposta terão uma melhoria considerável no tempo de análise dos dados, tornando a sua tomada de decisão mais eficiente e com mais segurança uma vez que vão conseguir ter um conjunto de factos para suportarem essa decisão.

A comparação dos cuidados de obstetrícia entre instituições torna-se uma mais valia para verificar a qualidade desses cuidados, a ferramenta proposta vai também fornecer essas comparações, podendo os profissionais tomarem decisões em conformidade com esses valores.

2.2.4 Modelo de negócio CANVAS

Para se obter uma visão global do modo como se pretende criar e entregar valor para o cliente normalmente recorre-se a representações estruturadas, como é o caso do modelo de negócio CANVAS.

Este modelo está estruturado em nove blocos que são:

- Segmentos de clientes;
- Proposta de valor;
- Canais;
- Relacionamento com clientes;
- Receitas;
- Recursos-chave;
- Parcerias-chave;

- Custos.

O modelo inicia-se pelo preenchimento do segmento de clientes identificando-se os clientes a quem se fornece valor. De seguida é elaborada a proposta de valor, o que se pretende oferecer a esses clientes.

Depois de se identificarem os clientes e o que se vai oferecer é necessário avaliar como se vai entregar esse valor aos clientes, identificando-se assim os canais de distribuição. Também é necessário perceber como se vai fortalecer os laços com esses clientes, como vai ser efetuado o relacionamento.

De seguida, é verificado como a solução vai gerar receitas e quais os recursos-chave que nos vão permitir chegar ao cliente. Nas atividades-chave identifica-se quais as atividades necessárias para se conseguir entregar valor ao cliente. Os parceiros-chave representam todos aqueles que nos vão poder ajudar na concretização desta solução. O bloco dos custos apresenta os custos necessários para o desenvolvimento e para a manutenção da solução.

A figura 2.1 apresenta o canvas com a identificação dos vários blocos para a solução que se vai desenvolver.

Parceiros-chave	Atividades-chave	Proposta de valor	Relacionamento com clientes	Segmentos de clientes
CINTESIS (Especialistas na análise de dados em saúde); ARSN, CHSJ (instituições onde inicialmente se vai instalar a aplicação, Beta testers) Comissão nacional Cesarianas (efetua a escolha dos indicadores) ISEP (recursos humanos para elaboração da ferramenta)	Lev. Requisitos para BI; Escolha e utilização de uma plataforma; Programação/Calculo dos indicadores, processo de ETL, geração de relatórios Instalação da solução Formação Manutenção	Apoio à decisão na gestão clínica de departamentos de obstetrícia hospitalares, através do fornecimento de um conjunto de indicadores; Possibilidade de comparar os indicadores de várias instituições de saúde.	Formação para a utilização da aplicação; Serviço de acompanhamento de utilização e parametrização	Hospitais públicos ou privados de países de língua oficial portuguesa com maternidades que tenham o ObsCare instalado; Administrações Regionais de Saúde de países de língua oficial portuguesa com maternidades que tenham o ObsCare instalado.
	Recursos-chave Recursos humanos (Programador; Tester, Prof. Saúde que valida os indicadores); Software (BD operacional com dados de obstetrícia, plataforma de BI, BD intermédia); Hardware (Servidores para: BD operacional; plataforma BI, BD intermédia)		Canais Contacto pessoal; Congressos científicos; Visitas aos hospitais; Manutenção à distância.	
Custos Deslocações		Receitas Serviços de instalação; Licença de utilização; Serviço de exploração de dados; Manutenção preventiva; Manutenção corretiva; Manutenção evolutiva; Apoio ao utilizador; Gestão de projeto		

Figura 2.1: Modelo de negócio, CANVAS.

2.3 Business Intelligence

É difícil encontrar uma definição global para o termo BI uma vez que se trata de um conjunto de processos. O conceito, de uma forma alargada, representa a recolha, armazenamento, processamento, análise e partilha de um conjunto de informações que fornecem suporte na gestão de um negócio. Este conceito surge em 1989 pelo analista Howard Dresner da Gartner Research (Nylund, 1999), onde define o termo BI como: "a broad category of software and solutions for gathering, consolidating, analysing and providing access to data in a way that lets enterprise users make better business decisions"(Gibson, Arnott e Jagielska, 2004)

Segundo, Turban, BI é um termo abrangente que possui várias ferramentas, arquiteturas, Base de Dados (BD), *Data Warehouse*, performance, gestão, metodologias integradas num único software. A junção de vários serviços e tecnologias para recolher, organizar e verificar os dados vai permitir converter os dados em informação funcional para levar a melhores julgamentos (Turban, Sharda e Delen, 2010).

Com o passar do tempo têm surgido várias definições englobando três categorias principais: processo, aspeto técnico e aspeto do produto. Do ponto de vista do processo, a definição concentra-se no processo de recolha dos dados das várias fontes externas e a sua análise de forma a gerar informações relevantes para melhorar o processo de tomada de decisão. O aspeto técnico foca-se nas ferramentas e tecnologias que permitem guardar, recuperar, manipular e analisar a informação. Outra forma de descrever o termo BI é sob a forma de produto em que a definição é voltada para o resultado de uma análise detalhada de um negócio, utilizando ferramentas de BI (Chee et al., 2009b).

A junção das várias ferramentas leva assim a um sistema de BI. Estes sistemas vão assim permitir às organizações tirar melhor partido de dados que tenham armazenados, não só atuais mas também histórico, oferecendo uma melhor gestão no planeamento do seu negócio. Para que o sistema forneça o apoio necessário à tomada de decisão é importante que possua uma arquitetura com uma estrutura que permita extrair dados de várias fontes, efetuar os tratamentos necessários, armazenar e aceder a estes. Este acesso pode ser efetuado de diferentes formas possibilitando ao utilizador visualizar os dados sob diversas formas, das quais se destacam os relatórios, *dashboards*, gráficos.

A figura 2.2 apresenta uma arquitetura típica de um sistema de BI com as várias estruturas associadas.

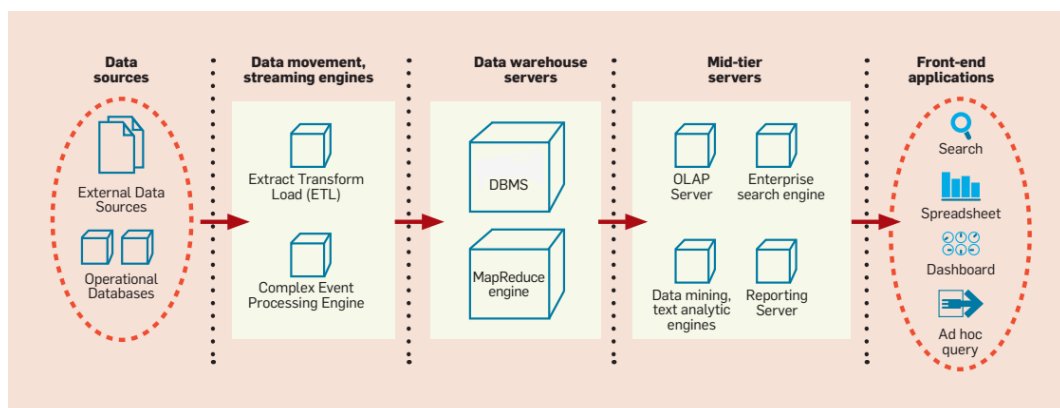


Figura 2.2: Arquitetura de um sistema de BI (Chaudhuri, Dayal e Narasayya, 2011).

Os dados que um sistema de BI recolhe normalmente vêm de diferentes fontes, que podem ser ficheiros de dados ou bases de dados operacionais, o que faz com que possuam representações inconsistentes e diferentes formatos. No entanto, estes dados têm de ser conciliados independentemente do formato em que se encontrem o que torna esta integração e standardização desafiante.

Para efetuar esta integração e preparação dos dados normalmente recorre-se a ferramentas de *Extract, Transform, Load* (ETL), numa base de dados específica (*Staging Area* (SA)). A existência da necessidade para suporte de tarefas de BI em "quase"tempo real, requer outro tipo de abordagem. Neste caso, o processo de tomada de decisão é baseado diretamente nos dados operacionais, utilizando ferramentas específicas para estes processamentos (Chaudhuri, Dayal e Narasayya, 2011). Depois dos dados estarem limpos e transformados, tipicamente, são guardados num *Data Warehouse* (DW). Nesta estrutura os dados são organizados de forma a agilizar as pesquisas e a disponibilidade da informação. A utilização de *MapReduce* é vulgarmente utilizado para analisar documentos *Web*.

Os dados são estruturados para que facilmente sejam atualizados e acedidos, no entanto, é necessário ter outras estruturas para efetuar as análises necessárias para os gestores de negócio. É o caso de servidores *Online Analytical Processing* (OLAP), processamento analítico online; servidores de relatórios; mecanismos de pesquisa; mecanismos de *data mining* (obtenção de padrões e previsão de modelos).

Para que o utilizador interaja com a informação de forma a obter o que pretende, são utilizadas ferramentas onde o utilizador pode visualizar toda a informação sob a forma de *dashboards* e relatórios onde poderá efetuar pesquisas mais aprofundadas com *ad-hoc queries*.

2.3.1 Processo de ETL

O processo de ETL está dividido em três partes, extração (*extraction*), transformação (*transformation*) e carregamento (*load*).

A extração representa o primeiro passo antes de efetuar a inserção dos dados no DW. Neste processo as diversas fontes externas são acedidas, sendo efetuada a leitura dos dados e de seguida copiados para o sistema de ETL, onde são efetuadas todas as manipulações necessárias.

Como os dados provêm de diferentes fontes é necessário efetuar várias transformações, entre elas, a limpeza de dados (correção de erros ortográficos, falta de elementos), a combinação dos vários dados e a verificação de duplicados. É importante nesta fase garantir a qualidade dos dados que serão usados pelo DW.

O carregamento dos dados é a fase em que a informação é colocada disponível para o utilizador.

2.3.2 Data Warehousing

Seguindo a metodologia de Kimball, *Data Warehousing* representa três fases, que são (Kimball e Ross, 2015):

- Ambiente Operacional;
- *Staging Area*;
- DW.

O ambiente operacional representa as várias fontes de dados operacionais, onde são inseridos os dados da organização. Nesta fase é importante que os dados estejam sempre disponíveis e acessíveis para que se possa efetuar a extração de todos os dados necessários.

A SA representa uma camada intermédia de armazenamento onde ocorrem os processos ETL. Esta área de armazenamento faz a ligação entre o ambiente operacional e o DW ou *Data Mart*, dependendo da arquitetura, e não é acessível pelos utilizadores.

O DW é a estrutura de dados que possuiu informação estruturada e consistente (dados credíveis). Esta estrutura tem de possuir a informação acessível, uma vez que possui a informação das várias fontes operacionais, bem como o histórico para auxiliarem os profissionais no processo de tomada de decisão.

Arquiteturas do DW

Devido ao grande volume de dados utilizados nos sistemas de BI é necessário ter arquiteturas específicas para o DW para que possa suportar estes volumes, de forma a que mesmo assim sejam facilmente acessíveis. As duas arquiteturas que se podem destacar são a de Inmon e de Kimball.

Bill Inmon conhecido como o pai do Data Warehousing, defende que "A data warehouse is a subject-oriented, integrated, Non-volatile and time-variant collection of data in support of management's decisions"(Inmon, 2002).

Para este autor o DW deve orientar-se às áreas de interesse da organização, a um assunto, por exemplo, o cliente. Refere também a uniformização do DW, guardando os dados das diferentes fontes no mesmo formato. A não volatilidade é devida ao facto de serem movimentados grandes volumes de dados de cada vez no DW, ao contrário do que acontece nos sistemas operacionais, em que os registos são inseridos um a um. Por último, Inmon defende que os dados estão associados a um dado momento. Para isso é necessário guardar histórico e guardar um valor que identifique a que momento correspondem aqueles dados.

Para este autor os dados provêm das diferentes fontes e vão assim popular um DW de onde irão derivar vários *Data Mart* (DM), tal como é apresentado na figura 2.3.



Figura 2.3: Arquitetura de Inmon.

Neste modelo os dados estão normalizados (possuem relações de forma a diminuir a redundância de dados) no DW, uma vez que é onde ocorre a integração das várias fontes, a desnormalização (criação de redundâncias para otimização dos resultados das consultas, uma vez que não é necessário efetuar o cruzamento de várias tabelas para obter os resultados) dos dados dá-se nos DM. Esta arquitetura é conhecida como *Hub and Spoke*.

A outra arquitetura é apresentada por Ralph Kimball, onde defende que o DW é a junção dos vários DM (Kimball e Ross, 2015). Para este autor o DW deve facilitar o acesso aos dados, tornando-se intuitivo para todos os utilizadores independente do seu nível técnico, para isso todos os dados têm de ter significado.

Os dados armazenados devem ser os corretos, e para isso os dados que provêm das fontes têm de ser analisados e limpos, caso necessário, sendo posteriormente apresentados aos utilizadores.

O DW deve ser flexível de forma a precaver possíveis alterações (nos dados das fontes) permitindo em simultâneo o armazenamento de um histórico de dados.

O acesso aos dados deve ser restrito como forma de garantir intrusões indevidas, protegendo assim os dados armazenados. O sucesso de DW não é obtido pelas técnicas de construção do modelo, mas sim pela aceitação e uso correto por parte dos utilizadores finais.

Esta arquitetura é conhecida como *Bus Architecture*, os dados provêm de diferentes fontes externas e populam o DW que é constituído pelos vários DM, das diferentes áreas de negócio, tal como é demonstrado na figura 2.4. Nestas estruturas os dados encontram-se desnormalizados (Kimball e Ross, 2015).

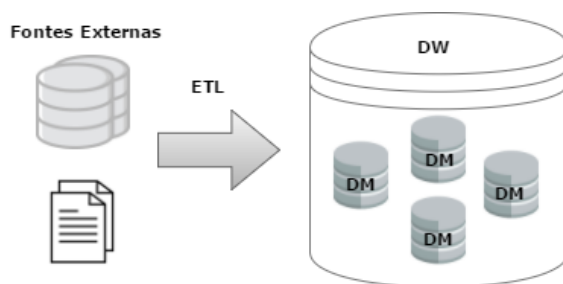


Figura 2.4: Arquitetura de Kimball.

De uma forma genérica pode-se dizer que o processo de ETL para Inmon é otimizado uma vez que possui os dados das diferentes fontes mas guardados num único formato uniformizando o DW. O DW possui todas as relações entre os dados, estando assim normalizado, minimizando as redundâncias o que leva a que o processo não transforme os mesmos dados

várias vezes. Para que o utilizador tenha acesso à informação é necessário derivar para um DM com uma modelação dimensional ¹. A manutenção de toda esta estrutura é complexa, sendo necessário efetuar um grande investimento inicial no desenvolvimento de toda a infraestrutura. O facto de ser complexa leva a que o seu acesso só possa ser efetuado por pessoas especializadas.

Para Kimball, o desenho do DW deve ser efetuado de forma a que todos os utilizadores tenham acesso e compreendam os dados, o que torna assim esta estrutura mais simples e perceptível (todos os dados têm significado). Este autor apresenta uma estrutura que não necessita de um custo elevado na sua implementação, o que pode ser um fator para determinar a escolha desta arquitetura ao invés de outra. O DW tem de permitir as várias mudanças que existem nos sistemas operacionais, devendo guardar a informação passada, mas não deve permitir a sua alteração (os dados devem permanecer inalterados). Os dados são constantemente extraídos, transformados e carregados para o DW.

A tabela 2.1 apresenta de forma resumida as características das duas arquiteturas.

Tabela 2.1: Características das arquiteturas de Inmon e Kimball.

	Inmon	Kimball
Processo ETL	Otimizado	Repetido
Orientação	Área de negócio	Processo de negócio
Complexidade	Elevada	Reduzida
Armazenamento de histórico	Guardar com data de inserção	Chaves primárias
Público	Pessoal especializado (IT)	Qualquer utilizador
Objetivo de desenvolvimento	Solução baseada em métodos	Solução de fácil utilização e compreensão
Investimento Inicial	Elevado	Reduzido

Constata-se que ambos os autores abordaram a construção de um DW de formas distintas, é necessário identificar-se claramente as necessidades da organização, bem como o tempo para a implementação e os recursos financeiros que têm disponíveis para se poder construir o sistema com a arquitetura que melhor se adequa.

2.4 Business Intelligence na Saúde

A maior parte da literatura referente a sistemas de BI é focada em organizações industriais, no entanto, o benefício destes sistemas já são amplamente conhecidos na saúde (Mettler e Vimarlund, 2007; Tremblay, Hevner e Berndt, 2012). Este sistema pode ajudar não só questões clínicas mas também administrativas (financeiras).

A indústria da saúde sofre uma elevada pressão para reduzir os custos e gerir os seus recursos de uma forma mais eficiente para melhorar constantemente os cuidados de saúde prestados (IBM, 2013). Isto leva a que as organizações tenham de fazer cada vez mais com menos, tendo de procurar formas de atingir os seus objetivos e gerir os recursos que têm disponíveis

¹Modelo desnormalizado constituído por dimensões e tabelas de factos (tabelas com as medidas).

da melhor forma (Hanson, 2011). O setor da saúde possui um grande volume de dados, estes devem ser analisados dando suporte no processo de tomada de decisão. Para um bom processo de decisão devem ser extraídos dados de várias fontes, por exemplo, sistemas de contabilidade e registos eletrónicos de saúde, sendo posteriormente disponibilizados aos médicos e outros utilizadores (Coddington e Moore, 2012).

Os sistemas de BI vão permitir efetuar toda a recolha dos dados das diferentes fontes, e aplicando ferramentas de análise vai fornecer as informações necessárias aos profissionais. Este tipo de sistemas permite, por exemplo, a identificação de tratamentos, programas e processos que sejam demasiado caros tomando-se depois medidas para os tornar mais eficientes ou mesmo substituí-los (IBM, 2013). Os profissionais vão assim desempenhar determinadas ações com base na informação fornecida, por exemplo, gestão de cuidados crónicos, a gestão dos profissionais de saúde e orçamentos (Coddington e Moore, 2012).

O aumento constante dos dados leva a que as organizações se voltem para estes sistemas, uma solução que para além de fornecer informações mais precisas para a tomada de decisão, vai permitir também reduzir custos e garantir o futuro da organização de saúde. O acesso a informações relevantes no momento oportuno é o primeiro passo para atingir estas metas (Ashrafi, Kelleher e Kuilboer, 2014).

A informação é a chave para um negócio de sucesso, e tal como noutras indústrias, pode facilmente aplicar-se um modelo simples de planear, fazer, verificar e agir (*Plan, Do, Check, Act*) (Ashrafi, Kelleher e Kuilboer, 2014). Todas estas fases devem ser executadas com informação o mais precisa possível, entrando aqui os sistemas de BI que fornecem esta informação. Estes sistemas podem também intervir na monitorização das medidas implementadas, verificando-se assim o seu impacto, se as medidas responderam da forma prevista ou não. Depois de efetuarem a verificação deste impacto os utilizadores terão de agir em conformidade com esses resultados, isto é, se as medidas aplicadas inicialmente não atingiram os objetivos pretendidos terá de se atuar de forma a encontrar novas soluções.

O grande desafio é converter um grande volume de dados em informação e conhecimento, para isso é necessário haver uma estrutura que suporte esta quantidade de dados com elevadas capacidades de performance. Para a indústria da saúde a capacidade de armazenar histórico é de extrema importância (Ashrafi, Kelleher e Kuilboer, 2014), no entanto, o mapeamento desse histórico através de diferentes fontes é difícil, centralizar os dados vai assim tornar mais fácil a disponibilização dos dados tornando mais simples o seu mapeamento.

Estes problemas são colmatados pela utilização de um DW, componente do sistema de BI. Este componente vai assim organizar, agrupar, classificar e estruturar informação que é benéfica para o utilizador, permitindo também efetuar consultas, conduzir pesquisas e analisar tendências. Neste setor o DW é essencial para que as decisões de negócio e clínicas sejam tomadas com base no maior conjunto de informação possível (Data et al., 2014).

Numa organização de saúde pode-se ter três tipos de fontes de dados: clínicos, financeiros e operacionais. As fontes de dados clínicos incluem todos os dados médicos (registos de pacientes, resultados de laboratórios, entre outros) que são necessários para a prestação de serviços de saúde aos pacientes. Dados financeiros incluem dados referentes à situação financeira dos vários departamentos; valor económico dos recursos para o hospital, bem como os gastos dos pacientes na organização. Os dados operacionais são os dados das operações, dos processos da organização, produtividade (Raghupathi e Raghupathi, 2013).

A utilização destes sistemas nos cuidados de saúde permitiu que os vários dados sejam fornecidos diretamente aos profissionais de saúde, efetuando assim o melhor uso destes. O processo de decisão tornou-se assim mais eficiente porque os utilizadores podem aceder a qualquer tipo de informação num tempo de resposta consistente e rápido, independentemente do volume de dados analisados (Chee et al., 2009a). Estes sistemas começaram por ser utilizados na parte administrativa mostrando assim os seus benefícios da utilização de um DW para gerir os hospitais (Evans, Lloyd e Pierce, 2012).

Os registos médicos dos pacientes são registados no registo médico eletrónico do paciente (*Electronic Medical Records* EMR), estes guardam assim um grande volume de dados que podem e devem ser armazenados em DW. Com este tipo de sistemas é possível para as instituições constituídas por várias unidades hospitalares obter informação de todas elas de uma forma mais simples e rápida.

Um exemplo deste sistema é implementado pela Intermountain HealthCare (Intermountain), sistema de saúde sem fins lucrativos que opera em 22 hospitais e em mais de 179 clínicas em Utah e Idaho (Evans, Lloyd e Pierce, 2012). Esta organização seguiu a metodologia de Kimball para definir a arquitetura de todo o sistema implementado. Os dados são recolhidos de várias fontes, que podem ser internas (diferentes departamentos) ou externas (bases de dados, Estado) e posteriormente armazenados num DW constituído pelos vários DM² que representam cada departamento, cirúrgico, financeiro, entre outros. Os dados armazenados são posteriormente analisados com cubos OLAP ou acedidos por outras ferramentas de relatórios, fornecendo assim aos utilizadores a informação necessária. A figura 2.5 apresenta a arquitetura conceptual da Intermountain HealthCare.

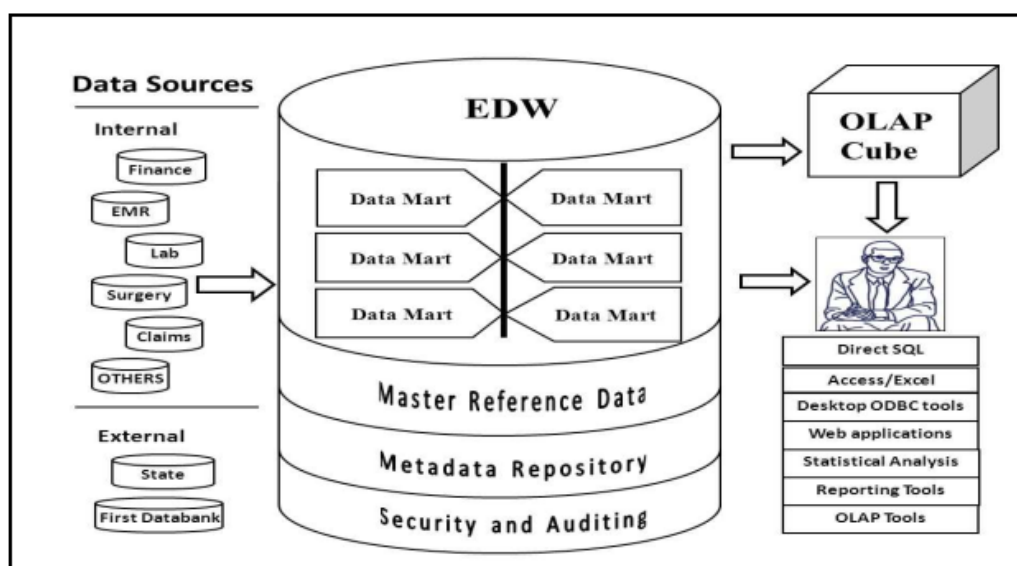


Figura 2.5: Arquitetura conceptual de um sistema de BI para o Intermountain HealthCare (Evans, Lloyd e Pierce, 2012).

Em Portugal, existe um sistema de consolidação da informação das unidades prestadoras de cuidados de saúde primários, sistemas regionais e centrais que é o Sistema Informação Monitorização do Serviço Nacional de Saúde (SIM@SNS). Este sistema estrutura os dados em três sistemas de BI (Serviços Partilhados do Ministério da Saúde, 2017):

²Bases de dados modeladas de forma multidimensional orientadas a departamentos das organizações.

- Sistema de Monitorização das Unidades Funcionais (MIM@UF) - Explora os dados das unidades funcionais dos cuidados de saúde primários;
- Sistema de Monitorização da Administração Regional de Saúde (ARS) - Exploração de dados a nível regional;
- Sistema de Informação e Monitorização do Serviço Nacional de Saúde - Exploração de dados a nível central, este sistema agrega os dados a nível nacional, servindo as necessidades de informação de organismos centrais, por exemplo, direção geral de saúde, administração central do sistema de saúde.

Esta plataforma disponibiliza um conjunto de indicadores e relatórios padronizados a nível nacional, regional ou local.

Apesar de existirem sistemas que fornecem informação generalizada, pode existir a necessidade de desenvolvimento de sistemas específicos, por exemplo, desenvolveu-se um sistema de BI para grupos de diagnóstico homogêneos (Barrento et al., 2010). A decisão por um sistema de BI deve-se pela necessidade de analisar de forma acessível e dinâmica a informação dos grupos de diagnóstico homogêneos, informação essa que se encontra dispersa por várias fontes. Para dar resposta aos utilizadores foi desenvolvido um sistema de BI, a arquitetura desse sistema encontra-se apresentada na figura 2.6.

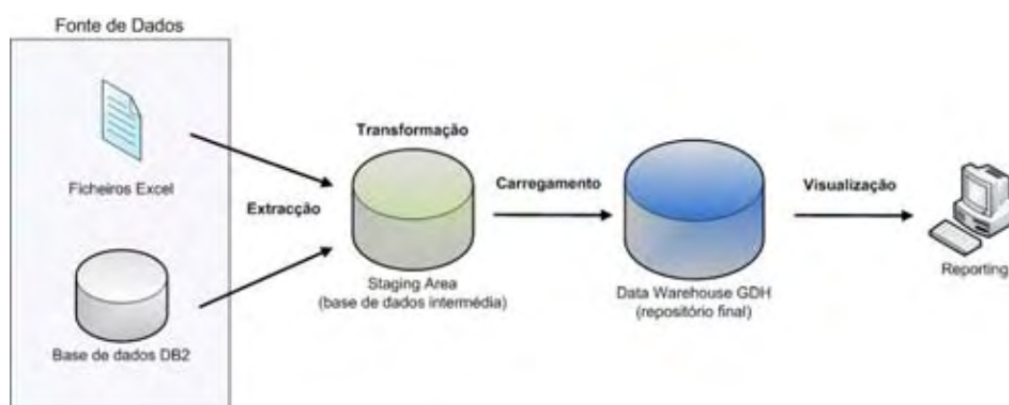


Figura 2.6: Arquitetura conceptual do sistema de BI aplicado ao grupo de diagnósticos homogêneos (Barrento et al., 2010).

Este sistema efetua a extração de duas fontes externas: ficheiros excel e BD (com os dados transacionais). Estes dados são transformados na SA e posteriormente armazenados num DW desenvolvido especificamente para este sistema. O modelo utilizado para o desenvolvimento do DW foi um esquema em estrela de forma a facilitar o acesso aos dados armazenados. A escolha por este sistema permitiu alargar as análises efetuadas, focando-se mais no paciente (Barrento et al., 2010).

2.5 Serviço de Obstetrícia

Alguns serviços têm indicadores específicos, padronizados, que é necessário fornecer a diferentes entidades, é o caso do serviço de obstetrícia que precisa de efetuar a monitorização dos seus próprios indicadores (Kao et al., 2016). Existem um conjunto de indicadores que é importante que sejam monitorizados, é o caso da taxa de mortalidade da mãe, do recém nascido (estes indicadores não são utilizados para índices de qualidade uma vez que ocorrem com menos frequência) e a taxa de cesarianas (Kao et al., 2016; Mann et al., 2006). Algumas instituições de saúde utilizam a taxa de cesarianas e a taxa de partos vaginais após uma cesariana para medirem a qualidade do serviço de obstetrícia (Mann et al., 2006). A taxa de cesarianas pode mesmo ser considerada como um indicador internacional para a qualidade dos cuidados obstétricos (WHO, UNFPA e UNICEF, 2009; Long et al., 2015).

Para a obtenção dos indicadores necessários neste serviço é necessário que ocorra uma boa recolha dos dados daquele serviço, sejam eles referentes às mulheres e aos seus bebés ou mesmo referentes aos recursos existentes neste setor.

Em Portugal esta recolha de dados é garantida pelo ObsCare, pelo menos nos hospitais pertencentes à ARS Norte. Este sistema efetua toda a recolha dos dados das pacientes, de todas as informações clínicas das suas gestações e dados dos respetivos nascimentos. Com todos estes dados é possível obter um conjunto de indicadores obstétricos, alguns destes indicadores encontram-se na 2.2.

Tabela 2.2: Indicadores do serviço de obstetrícia.

Indicador	Descrição
Nº de partos	Número de partos efetuados num determinado local, num determinado período de tempo
Nº de partos vaginais	Número de partos vaginais efetuados num determinado local, num determinado período de tempo
Nº de cesarianas	Número de cesarianas realizadas num determinado local, num determinado período de tempo
Nº de recém nascidos	Número total de recém nascidos num local num determinado período de tempo
Nº de recém nascidos vivos	Número total de recém nascidos vivos num local num determinado período de tempo
Nº de nados-mortos	Número total de nados-mortos num local num determinado período de tempo
Percentagem de cesarianas	Percentagem de partos que ocorreram por cesariana num número total de partos.

A DGS estabelece alguns indicadores para a monitorização da qualidade, sendo que são divididos em quatro categorias: Intervenções preventivas; Volume de utilização; Segurança; Pediátricos. Os indicadores por eles definidos para monitorizarem a qualidade para o serviço de obstetrícia encontram-se nas categorias: Volume e utilização; Segurança (Direção-Geral da Saúde, 2016).

Na categoria de volume e utilização os indicadores definidos são:

- Percentagem de partos por cesariana;
- Percentagem de 1º parto por cesariana;
- Taxa de partos vaginais após cesariana;
- Taxa de Histerectomia.

Na categoria de segurança os indicadores são:

- Trauma obstétrico em parto vaginal com instrumentação;
- Trauma obstétrico em parto vaginal sem instrumentação.

Para a DGS é crucial esta monitorização não só para perceber no momento como está o serviço de obstetrícia como também para avaliar a evolução da qualidade dos cuidados prestados ao longo do tempo. Outra entidade que precisa desta monitorização é o próprio hospital, os diretores de serviço, com o fornecimento dos vários indicadores e de outras informações eles podem tomar melhores decisões, decisões essas que podem ser efetuadas para uma melhoria nos cuidados de saúde prestados.

2.6 Conclusão

Após a elaboração do estado da arte é possível concluir a importância da implementação de um sistema de BI na saúde e as vantagens adjacentes. Vantagens estas que, pela capacidade de um rápido processamento e análise de um grande volume de dados, permitem ao utilizador ter acesso a informação de uma forma mais rápida e intuitiva. Desta forma é possível despende de mais tempo para a tomada de uma decisão clínica possibilitando a otimização de processos e a redução de custos, garantindo assim um aumento da qualidade para o serviço de saúde e para o utente.

Com o estudo efetuado foi possível perceber que a implementação de um sistema de BI nos serviços de obstetrícia é uma mais valia não só para os profissionais do hospital como também para entidades externas que necessitem dos dados para avaliações de qualidade e auditorias. Este tipo de sistema vai permitir-lhes um rápido processamento dos dados o que leva a que a disponibilização da informação também seja efetuada de uma forma mais rápida. Os sistemas de BI são sistemas de apoio à decisão, assim sendo, toda a informação disponibilizada pode ser usada pelos profissionais para suportarem as suas decisões. A disponibilização de diferentes indicadores, do serviço de obstetrícia, por este sistema vai ajudar os profissionais na monitorização do funcionamento do serviço.

Um sistema de BI vai permitir-lhes também efetuar o acompanhamento das variadas taxas, por exemplo a taxa de cesarianas, podendo assim, os diferentes utilizadores aplicarem a metodologia PDCA, referida na secção 2.4. Os utilizadores podem planejar como baixar a taxa de cesarianas, quais as medidas que vão ser tomadas e implementá-las. Depois de efetuarem o que planearam podem monitorizar o impacto dessas medidas, monitorização essa que pode ser efetuada pelo sistema de BI. Os utilizadores verificam se de facto as medidas que implementaram tiveram o resultado pretendido, a redução do número de partos por cesariana. Dependendo desse resultado os utilizadores podem agir em conformidade, definindo novas medidas ou mesmo prevalência das iniciais dependendo do resultado obtido.

Capítulo 3

Ferramentas de Business Intelligence

Neste capítulo é apresentada uma análise das ferramentas de análise de dados existentes com mais destaque no mercado e quais as suas características. Das ferramentas encontradas foi efetuada uma escolha com base no método AHP, sendo apresentados também os diferentes critérios considerados na avaliação.

3.1 Análise das ferramentas

Uma vez que os dados se encontram armazenados é necessário analisá-los e disponibilizá-los para os utilizadores. Atualmente, existem diversas tecnologias para se poder efetuar essa análise, algumas é necessário despende de recursos económicos para a aquisição de licenças, outras não apresentam qualquer custo e outras disponibilizam o código mas pode ser necessário pagar para adquirir a própria ferramenta. Todos os anos a Gartner, organização de consultoria e pesquisa em tecnologias de informação, efetua pesquisas e apresenta análises sobre empresas e as suas comparações para ajudar as organizações a optarem pela melhor ferramenta (Gartner, 2017). Para efetuar estas avaliações e comparações, a Gartner define 14 capacidades críticas de uma plataforma de BI e de análise. Estas características encontram-se divididas por quatro categorias: Infraestrutura; Gestão dos dados; Análise e criação de conteúdo; Partilha de resultados (Parenteau et al., 2016).

Na categoria de infraestrutura as ferramentas têm de possuir quatro características:

- Administração da plataforma de BI - Deve possuir capacidades para dimensionar a ferramenta, otimizando o desempenho e garantindo a disponibilidade e recuperações em caso de erros;
- *Cloud BI* - A plataforma pode estar alojada na *cloud* e ser utilizada como um serviço (SaaS - *software as a service*) permitindo de igual modo executar uma análise aos dados armazenados;
- Segurança e administração de utilizadores - Funcionalidade que tem como principal intuito preservar a segurança dos dados e da plataforma através da gestão de utilizadores. Deve ainda ser incluída a possibilidade de agendar e realizar auditorias ao sistema;
- Conectividade com as fontes de dados - Capacidades que permitam aos utilizadores conectar com dados estruturados e não estruturados contidos em diferentes tipos de plataformas de armazenamento.

Para gestão dos dados existem três características:

- Gestão de metadados - As ferramentas devem fornecer uma forma robusta e centralizada para os administradores poderem pesquisar, capturar, armazenar, reutilizar e publicar objetos de metadados, por exemplo, dimensões, hierarquias, medidas, métricas de desempenho;
- Incorporação do ETL e armazenamento - A ferramenta tem de ter a capacidade de aceder, integrar, transformar e carregar os dados para uma camada interna específica. Deve ter a habilidade para efetuar indexação, gestão dos carregamentos bem como agendamentos de *refresh* aos dados;
- Preparação dos dados pelo utilizador - A ferramenta deve possibilitar a combinação de dados pelo utilizador, podendo criar modelos analíticos, definir medidas, grupos ou hierarquias.

Na categoria de análise e criação de conteúdo existem quatro características, que são:

- Análise avançada incorporada - Permitir aos utilizadores o acesso a capacidades avançadas de análise incorporadas na ferramenta ou disponíveis através de integrações com modelos externos;
- *Dashboards* analíticos - Possibilidade de criar *dashboards* altamente interativos, com exploração visual;
- Exploração visual interativa - Permitir a exploração dos dados através da manipulação de gráficos, entre outros. Estas ferramentas devem permitir ao utilizador analisar os dados interagindo diretamente com a sua representação visual;
- Exploração móvel - As ferramentas devem permitir às organizações desenvolver conteúdo para dispositivos móveis num modo de publicação e/ou interativo. Podem também tirar partido de ferramentas dos dispositivos móveis, como por exemplo, a câmara, o teclado, localização.

Por último, tem-se a categoria de partilha de resultados, onde é possível destacar, três características:

- Incorporação de conteúdo analítico - Funcionalidades que incluam um kit de desenvolvimento de software e suporte para *standards* abertos de criação e modificação de conteúdo analítico, visualizações e aplicações. Incorporando-os no processo de negócio. Estes recursos podem residir fora da aplicação (reutilizando a infraestrutura analítica) mas tem de ser facilmente acessível a partir do interior da aplicação sem que os utilizadores precisem de trocar de sistemas. A arquitetura da aplicação permitirá que os utilizadores possam escolher onde o processo de análise deve ser incorporado;
- Publicação do conteúdo analítico - Funcionalidade que permite aos utilizadores publicar, desenvolver, e operacionalizar conteúdo através diferentes *outputs* e métodos de distribuição;
- Colaboração e BI Social - Permite que os utilizadores partilhem e discutam informação, análises, conteúdo analítico e decisões através de anotações ou *chats*.

Consoante todas estas características é elaborado um quadrante mágico, apresentado na figura 3.1, em que anualmente as ferramentas são classificadas como desafiantes, líderes, nicho de mercado, visionárias.



Figura 3.1: Quadrante Mágico (Parenteau et al., 2016).

Inicialmente o requisito principal para a elaboração da solução de apoio aos profissionais de saúde era a utilização de uma ferramenta *open source*, no entanto, ao longo da elaboração desta dissertação verificou-se que se poderia optar por uma que não fosse *open source* desde que a sua relação custo benefício assim o justificasse. Avaliando o quadrante de Gartner verifica-se que existem empresas que fornecem versões *open source*, como a Pentaho e TIBCO software, e outra que apresenta uma solução com uma versão mais básica sem custos, e uma versão mais completa com custos, a Microsoft. Apesar de não se encontrarem contempladas no quadrante de Gartner existem outras ferramentas *open source* relevantes de outros fornecedores que merecem algum destaque, é o caso da Birt e SpagoBI (Predictive Analytics Today, 2017).

Verificou-se que estas cinco ferramentas (Pentaho, TIBCO, Microsoft, Birt, SpagoBI) se encontram com grande destaque no mercado, desta forma, foram analisadas com mais detalhe nas secções seguintes.

3.1.1 Pentaho

Pentaho é uma subsidiária de Hitachi Data Systems, possuindo uma plataforma para BI que fornece suporte a relatórios, transformação dos dados, capacidades analíticas avançadas e visualizações interativas da análise de negócio. Esta empresa oferece uma versão *open source*, *Community Edition* que inclui ETL, *Online Analytical Processing*, *data mining*, relatórios e *dashboards*.

Esta plataforma é constituída por uma variedade de soluções *open source* para fornecer as diferentes funcionalidades, por exemplo, a utilização de *Weka*¹ para *data mining*. A ferramenta pode efetuar a conectividade com as fontes de dados com um armazenamento de dados e um processo de ETL autónomo, permite também trabalhar paralelamente com outras ferramentas desenvolvidas em Java, possibilitando uma personalização de acordo com as necessidades do utilizador.

¹Coleção de algoritmos de *Machine Learning*

Caso a organização pretenda utilizar a *cloud*, a plataforma permite essa sincronização, o que pode aumentar significativamente a eficiência na análise de dados bem como a rapidez na distribuição para os utilizadores necessários. Os relatórios obtidos podem ser exportados em vários formatos, entre os quais PDF, Excel e CSV.

A Pentaho encontra-se posicionada no quadrante dos visionários devido a um posicionamento inovador no setor de *Big Data* e no espaço analítico da *Internet of Things* (Internet das coisas).

3.1.2 TIBCO software

Esta empresa foi um dos primeiros líderes na descoberta de dados, no entanto, continuou a expandir as suas ofertas para incluírem análises avançadas em tempo real e localização inteligente. Encontra-se no quadrante dos visionários devido à diversidade de capacidades que oferece.

A TIBCO adquiriu a Jaspersoft em 2014 para impulsionar as suas vendas de baixo custo. A Jaspersoft tem uma ferramenta de BI *open source*, a Jaspersoft Community Edition. A ferramenta pode ser integrada com qualquer aplicação ou dispositivo móvel, estando também disponível para utilização na *cloud*.

3.1.3 Microsoft

A Microsoft oferece uma ampla gama de recursos de BI e funcionalidades analíticas, o quadrante de Gartner, foca-se no Power BI que oferece soluções baseadas na *cloud*. A Microsoft encontra-se posicionada no quadrante dos líderes, uma vez que teve uma forte aceitação no mercado, apresentando também uma boa complexidade de visão com o fornecimento de consultas e exploração de dados com linguagem natural.

Esta ferramenta oferece transformação dos dados, descoberta dos dados e *dashboards* interativas. A versão 2.x permite a utilização de fontes de dados locais, não estando obrigatoriamente na *cloud*. A Microsoft apresenta uma versão gratuita, tendo um limite no volume de dados utilizado, sendo que a versão superior, também com um limite no volume de dados de 10GB, apresenta um custo de 8,40 euros por utilizador, por mês (Microsoft, 2017d), o que torna a ferramenta uma das soluções de menor preço no mercado dos grandes fornecedores (Parenteau et al., 2016). Algumas das funcionalidades apresentadas pela ferramenta são:

- Pré-instalação de uma instância local de R;
- Conexão com diferentes fontes de dados (Access, PostgreSQL, Hadoop, SQL Server, Oracle);
- Junção de várias ferramentas: Power Query (para extração e transformação dos dados); Power Pivot (para modelação e análise dos dados); Power View (para mapear e visualizar os dados);
- Atualização dos dados de forma automática;
- Exploração de dados através de linguagem natural (torna mais simples a pesquisa por informação);

- Criação de relações e alterações do modelo de dados para criação de novas medidas e formatos de dados;
- Análise dos dados em memória;
- Criação de medidas e colunas com base em scripts DAX (*Data Analysis Expressions*);
- Criação de relatórios. Permite selecionar os campos que se pretende, adicionar filtros, escolher o tipo de visualização, escolher cores (por exemplo) para personalizar os relatórios;
- Upload e publicação dos relatórios. É possível publicar os relatórios para os utilizadores acederem, podendo estes interagir com os dados do seu lado.
- Configurar a atualização dos dados publicados de forma a que todos os utilizadores acessem à mesma informação e à mais atualizada.

3.1.4 Birt

Software *open source* para visualização de dados e relatórios, que podem ser incorporados em aplicações cliente e *Web* baseados em Java. Esta faz parte de um projeto de software da fundação Eclipse (consórcio independente sem fins lucrativos de fornecedores de software e uma comunidade *open source*) e é patrocinada pela Actuate, com contribuições da IBM e Innovent Solutions. A aplicação permite assim a incorporação de relatórios e *dashboards*, permite a visualização dos dados sob a forma de gráficos e tabelas. Os relatórios podem ser guardados em XML o que permite que sejam integrados posteriormente com outras aplicações, caso necessário (The Eclipse Foundation, 2014).

3.1.5 SpagoBI

Ferramenta *open source* que abrange todas as áreas analíticas de projetos de BI com motores e temas inovadores. SpagoBI oferece um conjunto de ferramentas analíticas que são (SpagoBI, 2017):

- Relatórios - Estruturação de relatórios e posterior exportação para diferentes formatos (HTML, PDF, XLS, XML, TXT, CSV, RTF);
- Análises multidimensionais (OLAP) - Explora os dados em diferentes níveis de detalhe e diferentes perspetivas, através de *drill-down*, *drill-across*, *slice-and-dice*, *drill-processes*;
- Gráficos - Desenvolvimento de gráficos prontos a utilizar, por exemplo, histogramas, diagramas de dispersão, entre outros e também gráficos interativos;
- Indicadores de performance - Possui um conjunto de ferramentas para criar, gerir, ver e navegar nos modelos de hierarquia dos indicadores através de métodos, regras de cálculo, entre outros;
- Cockpit interativos - Agrega diferentes análises numa única vista, define caminhos de visualização e explora os dados de forma dinâmica;
- Relatórios *Ad-hoc* - O utilizador pode criar os seus relatórios com várias folhas, tabelas, gráficos;

- Localização inteligente - Visualização dos dados em mapas, permitindo interação dinâmica para obtenção de vistas instantâneas;
- Consulta facilitada - A ferramenta torna a exploração dos dados particularmente intuitiva e simples, guardando as *queries* efetuadas para posterior reutilização;
- *Data mining* - Análises avançadas de dados que permitem a extração de conhecimento de grandes volumes de dados;
- Análises em rede - Visualização e interpretação de relações entre entidades através de vistas especializadas;
- ETL - A ferramenta integra outro produto *open source* Talend Open Studio, que permite carregar os dados para um DW e efetuar a sua gestão;
- Colaboração - Permite a criação de dossier de relatórios estruturados, permitindo colocar notas ou comentários pessoais. Posteriormente é possível partilhá-los para uma rede de trabalho colaborativa;
- Automação - Publicação dos documentos pessoais num ambiente de BI integrando com outras ferramentas (Open Office ou Microsoft Office);
- Gestão de dados - Os utilizadores podem escrever na base de dados e modificar dados através de uma interface que se comporta consoante parâmetros de configuração pré definidos;
- Processos Externos - Gere os processos analíticos que podem correr em *background* ou agendados para começar e parar num período de tempo definido.

3.2 Comparação das ferramentas

Os principais requisitos para um sistema de BI prendem-se com funcionalidades chave: o processo de obtenção e transformação dos dados (ETL), peça importante uma vez que sem eles não é possível efetuar nenhuma análise pelo utilizador; o processo de análise, onde é possível ao utilizador obter a informação necessária para obtenção de conclusões e geração de conhecimento; o processo de disponibilização da análise efetuada que pode ser, por exemplo, por relatórios ou *dashboards* e o custo que essa ferramenta terá para o utilizador.

Com o estudo efetuado verificou-se que as ferramentas encontradas apresentam as funcionalidades pretendidas para a solução que se pretende desenvolver, no entanto, utilizam diferentes métodos para chegarem a essas funcionalidades.

Como forma de se perceber as diferentes características de cada ferramenta em cada uma das funcionalidades, foi elaborada uma tabela de comparação 3.1.

Tabela 3.1: Verificação das características das ferramentas encontradas.

	Pentaho	Jaspersoft	Birt	SpagoBI	Power BI
ETL	Autónomo; Escalável; Ambiente gráfico (<i>drag and drop</i>)	Fácil de implementar; Integra diferentes fontes; Ambiente <i>drag and drop</i>	<i>Web Services</i> para aceder aos dados; Conecta com XML e objetos Java em memória	Integra outra ferramenta Talend Open Studio (ambiente gráfico <i>drag and drop</i> , facilmente escalável)	Integração com diferentes fontes. Utilização do Power Query.
Análise	Ferramentas de <i>Machine Learning</i> e <i>data mining</i> Servidor OLAP (Mondrian)	Servidor OLAP Processamento configurável para acesso em memória	Permite a utilização de scripts Java e JavaScript; <i>data mining</i> (<i>clustering</i> e <i>forecasting</i>); Ferramentas de segmentação (seleção)	OLAP (exploração em diferentes níveis de detalhe); <i>data mining</i> ; Criação e Gestão de KPI's (<i>Key Performance Indicators</i>)	Power Pivot e utilização de linguagem natural. Análise de dados em memória
Relatórios	Criação e partilha; Geração de relatórios em PDF, Excel, HTML, Txt, RTF, XML, CSV	Gestão de relatórios para acesso de utilizadores específicos	Conexão com a maioria das BD; Ambiente <i>drag and drop</i> para a criação do relatório	Criação de relatórios em HTML, PDF, XLS, XML, Txt, CSV, RTF	Personalização. Publicação dos relatórios e configuração do momento da atualização dos relatórios.
Dashboards	Utiliza AJAX; Separa a lógica (JavaScript) da apresentação (HTML, CSS); Extensível (elevado nível de personalização); Interativos	Disponível na versão <i>reporting</i> , parâmetros para controlar dados internos e externos	Interativos, possibilidade de interagir com os dados baixando ou subindo na granularidade; Ambiente <i>drag and drop</i> para a exploração dos dados	Exploração dos dados de forma dinâmica e gráfica	Power View, <i>dashboards</i> personalizáveis e interativas. O <i>dashboard</i> pode ser gerado através de mais que uma fonte (relatórios).
Custo	<i>open source</i> ; Versão base <i>free</i> , funcionalidades adicionais necessário comprar licença	<i>open source</i> ; versão <i>community, free</i> , licença AGPL, versão <i>reporting</i> , licença comercial paga por CPU	<i>open source</i> ; licença pública do Eclipse	<i>open source</i> e <i>free</i>	Power BI Desktop gratuita, versão Pro 8,40 euros por utilizador, por mês

De uma forma geral, todas apresentam as mesmas características, apesar de terem diferentes formas de atingir essas funcionalidade, é o caso de SpagoBI que comparativamente com as restantes, para garantir a funcionalidade de ETL utiliza uma ferramenta externa. A forma como se vão diferenciar será nas capacidades de adaptar essas ferramentas com outras ferramentas externas, outros sistemas, a escalabilidade a capacidade de configurar estas ferramentas de análise. Outra diferença também muito importante prende-se com o tempo, a capacidade de execução que estas ferramentas terão para realizar as funcionalidades que pretendemos.

3.2.1 Grandezas a Avaliar

Para se efetuar a escolha da ferramenta têm de ser definidos alguns critérios/funcionalidades que estas têm de apresentar. No entanto, tendo em conta as necessidades do cliente estas funcionalidades têm diferentes níveis de importância para ele. De forma a tornar esta escolha o mais imparcial e o mais rigorosa possível, será utilizado um método matemático, onde se terá de indicar as funcionalidades que terão de ser avaliadas e o peso que estas terão para o cliente.

Um método muito utilizado nestes processos de tomada de decisão é o Processo Analítico de Hierarquia (AHP) (Vaidya e Kumar, 2006). Esta metodologia atrai assim a atenção dos investigadores devido às suas propriedades matemáticas e à facilidade de aquisição dos parâmetros que o método necessita para auxiliar na escolha da melhor ferramenta.

AHP é uma ferramenta com uma estrutura hierárquica de vários níveis, que são: os objetivos, critérios e alternativas. As alternativas possíveis são comparadas duas a duas para cada critério, com base na avaliação em todos os critérios o método indica qual a melhor opção para o objetivo definido inicialmente. É um método quantitativo que indica qual a melhor escolha, alternativa com melhor pontuação, que se adequa ao utilizador e às suas prioridades. O método AHP é um método multicritério que pode ser aplicado para ajudar na escolha, na avaliação, com base em critérios pré-definidos.

Para esta dissertação este método vai permitir ajudar na escolha da ferramenta a ser utilizada para a análise dos dados de obstetrícia. Sendo o objetivo a escolha desta ferramenta, qual a ferramenta que melhor se adequa na análise dos dados. Face a este objetivo foram definidos critérios para comparação das ferramentas, critérios esses que foram definidos em conjunto com o utilizador.

Os critérios definidos são:

- Escalabilidade (capacidade da ferramenta de crescer para diferentes módulos, mais funcionalidades);
- Configuração (flexibilidade de parametrização da ferramenta);
- Performance (desempenho da ferramenta);
- Integração (incorporação com novos sistemas).

Para cada critério é atribuído um peso que se encontra diretamente relacionado com a importância do mesmo para o utilizador. Se para o utilizador o critério Configuração, por exemplo, for o mais importante de todos então terá um peso maior comparativamente aos outros. Todos os pesos serão definidos em conjunto com o cliente atribuindo-se assim um maior peso aos critérios com maior preferência.

Depois de definidos os critérios são indicadas as várias alternativas disponíveis, no caso desta dissertação, as alternativas encontradas são:

- Pentaho;
- Jaspersoft;
- Birt;
- SpagoBI;
- Power BI.

A figura 3.2 apresenta o problema, os vários critérios e as várias alternativas encontradas para solucionar o problema.

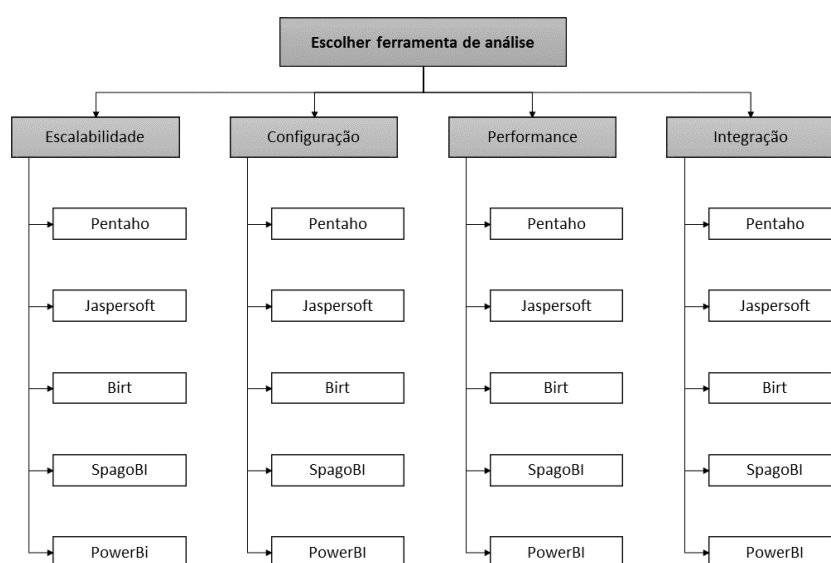


Figura 3.2: Esquematização do problema, dos critérios e das alternativas.

Para a construção do modelo foi necessário comparar os critérios dois a dois, verificando-se qual seria mais importante. Esta avaliação foi efetuada utilizando a escala de Saaty (Saaty, 1990), atribuindo-se valores de 1 a 9 em que 1 indica que os dois critérios têm a mesma importância, aumentando assim a grau de importância até ao 9. O valor máximo, o 9, corresponde a uma importância absoluta, um critério é favorecido com o mais alto grau de certeza. Esta comparação aos pares é utilizada para se construir uma matriz com a importância relativa entre eles. A matriz correspondente a esta comparação encontra-se na tabela 3.2.

Tabela 3.2: Comparação dos pesos de importância.

	Escalabilidade	Configuração	Performance	Integração	Prioridade relativa
Escalabilidade	1	$\frac{1}{2}$	$\frac{1}{5}$	$\frac{1}{3}$	0,0761
Configuração	2	1	$\frac{1}{5}$	$\frac{1}{3}$	0,1068
Performance	5	5	1	5	0,6086
Integração	3	3	$\frac{1}{5}$	1	0,2085

O peso de cada critério, tal como já foi referido, foi definido em conjunto com o cliente. De forma unânime percebeu-se que a Performance é a que possui um peso superior, uma vez que uma baixa performance pode levar a uma não aceitação da solução por parte do utilizador final. Se a solução proposta não apresentar as análises rapidamente aos profissionais de saúde pode levar a que estes deixem de utilizar a aplicação. Deste modo, este critério assume um peso no valor de 5 (importância grande) comparativamente com os outros.

A solução proposta necessitará de efetuar o carregamento de dados, estes encontram-se descentralizados, dispersos por diversas fontes. A ferramenta terá de aceder aos dados independentemente do tipo de sistema onde se encontram, deverá permitir integrar-se com uma diversidade de fontes e garantir uma boa conexão para que a disponibilização dos dados não se encontre limitada. Como tal, foi atribuído um peso de 3, comparativamente com a escalabilidade e a integração demonstrando-se assim que este critério é um pouco mais relevante que estes dois últimos

A configuração da ferramenta também é relevante, apresentando uma importância ligeiramente superior comparativamente com a escalabilidade (valor 2). Este critério é importante para que a ferramenta possa ser configurável e assim adaptar-se às preferências dos utilizadores.

A escalabilidade também não deixa de ser um critério importante apesar de possuir o valor mais baixo de prioridade. A escalabilidade na ferramenta irá permitir a adaptação de novos módulos futuros, por exemplo, a exploração de indicadores que não teriam sido contemplados na fase inicial.

Depois de definida a importância de cada critério é verificado em cada ferramenta os critérios definidos. As ferramentas são comparadas, no mesmo critério aos pares, construindo-se assim uma matriz com os valores obtidos. As várias matrizes de comparações dos diferentes critérios encontram-se no anexo A.

Com base nos resultados das diferentes comparações, conclui-se que o Power BI é a ferramenta que melhor se adequa ao problema, uma vez que é a que possui um valor mais alto. Todas as ferramentas foram avaliadas nos vários critérios e comparadas, o Power BI destacou-se em todos os critérios, exceto na Performance. Neste critério o Jaspersoft apresentou uma Performance semelhante ao Power BI. No entanto, na avaliação dos quatro critérios foi o Power BI que apresentou melhores resultados, sendo o escolhido para utilização na solução desenvolvida.

3.3 Conclusão

Através da pesquisa bibliográfica elaborada foi possível verificar as ferramentas mais relevantes no mercado, e quais apresentavam as melhores características para a resolução do problema encontrado. Com esta análise destacaram-se cinco ferramentas, que posteriormente foram analisadas com mais detalhe. As cinco ferramentas de BI foram avaliadas consoante um conjunto de critérios, para se poder efetuar a escolha da ferramenta a utilizar.

Para esta avaliação foi utilizado um método matemático, AHP, método baseado em hierarquias. Através deste verificou-se que a melhor ferramenta para o problema encontrado ao longo da dissertação foi o Power BI da Microsoft.

O Power BI apresenta elevada capacidade de processamento destacando-se pela sua performance nas respetivas análises. Para além disto é uma ferramenta muito intuitiva que apresenta um conjunto de funções para auxiliarem o utilizador nas análises pretendidas. Posto isto, será esta a ferramenta utilizada para analisar e disponibilizar os dados aos utilizadores.

Capítulo 4

Design da solução proposta

Neste capítulo são propostas duas arquiteturas para o problema encontrado. Destas arquiteturas é verificada qual responde melhor às necessidades procedendo-se à sua implementação. A implementação da arquitetura final é descrita também neste capítulo.

4.1 Introdução

Antes de desenvolver um sistema de BI é necessário definir cuidadosamente os processos, atores e tecnologias que vão ser utilizadas (Mettler e Vimarlund, 2007). Os processos podem ser definidos como um conjunto de tarefas, no contexto de BI, podem ser vistos como a forma de recolha dos dados, o processamento e a distribuição, podem também ser definidos como as práticas de trabalho dentro da organização de saúde, focados principalmente na prestação de serviços de saúde. Duas características importantes que constituem os processos: os processos têm clientes, ou seja os processos têm resultados definidos e respetivos destinatários; os clientes são atores que podem ser internos ou externos à organização de saúde.

O setor de saúde envolve um conjunto de atores com necessidades distintas (pacientes, companhias de seguro, médicos, autoridades governamentais), estes atores podem ser internos ou externos. Atores internos representam o pessoal interno da organização de saúde, por exemplo, os médicos, a quem tem de se fornecer um serviço personalizado e o acesso a informações atualizadas. Atores externos correspondem aos *stakeholders* que têm uma grande influência na organização de saúde (companhias de seguros, fornecedores, autoridades governamentais) mas são apenas incorporados parcialmente nos seus processos de informação, tendo na maior parte das vezes acesso restrito ao sistema.

Para fornecer um bom serviço de saúde deve conhecer-se os requisitos dos atores, no entanto, quando se desenvolve um sistema de BI, a heterogeneidade de atores pode causar problemas, uma vez que a mesma informação pode ser visualizada de forma diferente por estes, posto isto, é crucial envolver o máximo de atores quando se está a desenhar o sistema (Mettler e Vimarlund, 2007).

Para o problema identificado ao longo desta dissertação verificou-se que os atores que devem intervir no processo da elaboração da ferramenta são os profissionais de saúde, médicos obstetras e os profissionais da ARS. Tanto os médicos como os profissionais da ARS são atores internos, vão ser os utilizadores da aplicação e como tal, é necessário verificar os seus requisitos para que a ferramenta vá de encontro ao que pretendem.

Para este problema foram identificados pelo menos dois processos, independentemente da estrutura dos dados das fontes externas, que são:

- Extração, transformação e carregamento dos dados;
- Visualização/Disponibilização.

4.2 Requisitos

O levantamento de requisitos é uma atividade importante no desenvolvimento de um projeto, uma vez que é nesta etapa que se verifica o que o sistema terá de alcançar. Depois de verificadas todas as necessidades e quais a que o sistema irá de encontro é desenhada uma arquitetura para responder aos requisitos encontrados. Os requisitos podem ser divididos em requisitos funcionais e não funcionais.

4.2.1 Requisitos não funcionais

Os requisitos não funcionais representam fatores de qualidade não só para garantir as funcionalidades, mas também o bom funcionamento do sistema, para a ferramenta elaborada os requisitos são:

- Desempenho;
- Usabilidade;
- Histórico;
- Fiabilidade;
- Segurança.

A aplicação vai permitir ao utilizador gerar relatórios em tempo real, e como tal é necessário garantir que essa execução ocorra o mais rápida possível. O sistema tem de ser desenhado de forma a que o utilizador facilmente o consiga manipular obtendo, assim, tudo o que necessita. Caso esta utilização não seja clara e facilmente perceptível pode levar à não utilização do sistema desenvolvido.

As aplicações de BI caracterizam-se pela capacidade de fornecimento de histórico de dados, podendo assim fornecer dados anteriores para auxiliarem os utilizadores nas suas tomadas de decisão. Como estas aplicações transformam os dados em informação e posteriormente em conhecimento, é necessário garantir a fiabilidade dos dados para que a informação fornecida seja a mais correta possível.

A aplicação desenvolvida será para a área da saúde, em que a segurança de dados é extremamente importante. Para garantir a segurança é necessário que sejam cumpridos três fatores: disponibilidade dos dados (deve garantir-se que a informação esteja disponível para o utilizador e livre de qualquer erro); integridade dos dados (quando ocorrem alterações aos dados nas fontes deve-se garantir que essas alterações também são asseguradas pela solução, garantindo a exatidão e fiabilidade dos dados); confidencialidade (os dados externos tratam-se de informações clínicas de utentes, como tal, deverão ser anonimizados).

4.2.2 Requisitos funcionais

Os requisitos funcionais representam o que o sistema deverá fazer, ou seja, as suas funcionalidades. A ferramenta poderá ser utilizada por dois utilizadores distintos, os médicos dos hospitais ou profissionais da ARS, existem requisitos que são comuns para os dois utilizadores que são:

- Filtrar informação;
- Executar relatórios;
- Visualizar *dashboards*;
- Comparações interinstitucionais.

O utilizador ao utilizar o sistema deve poder filtrar informação, isto é, visualizar apenas a informação que pretende, por exemplo, visualizar um indicador para um determinado ano. O sistema deverá permitir que o utilizador possa executar os relatórios de forma a guardá-los em diferentes formatos. Para que o utilizador possa interagir com a informação o sistema deverá fornecer *dashboards*, onde o utilizador deverá visualizar a informação e interagir com esta.

Existe outro requisito específico para os profissionais da ARS que é a comparação de informação dos diferentes hospitais. Na figura 4.1 estão apresentados os vários casos de uso, para as várias interações do utilizador com o sistema.

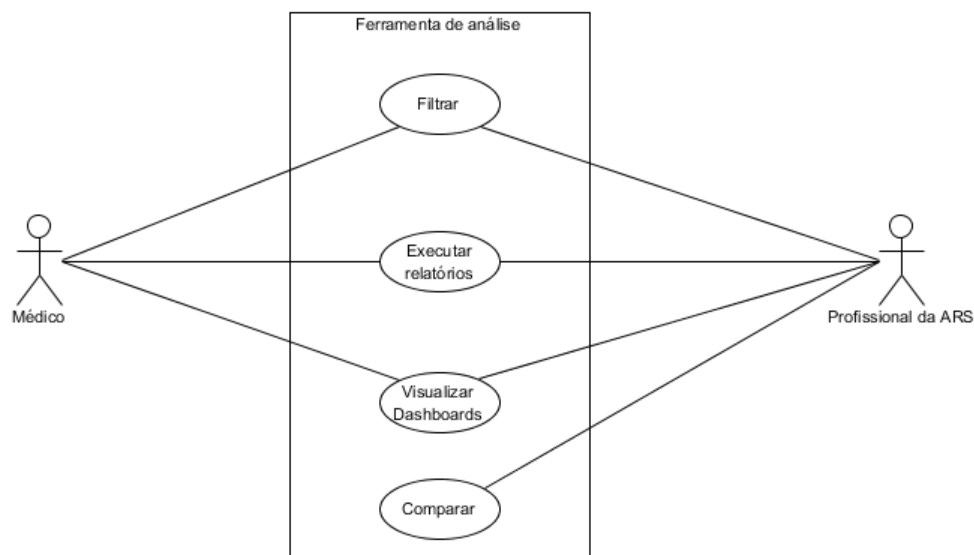


Figura 4.1: Diagrama de casos de uso.

Caso de uso 1: Filtrar informação

Os atores envolvidos neste caso de uso são o médico e o profissional de saúde.

Este caso de uso permite ao utilizador filtrar a informação que pretende visualizar com base em diferentes variáveis (idade da grávida, tipo de parto, semanas de gestação). Deste modo,

ele poderá acrescentar ou retirar informação à interface de visualização, consoante o filtro aplicado.

Caso de uso 3: Executar relatórios

Os atores são o médico e o profissional da ARS, o utilizador poderá aceder à ferramenta e executar relatórios com a informação que pretende analisar.

Caso de uso 4: Visualizar dashboards

Neste caso de uso estão também envolvidos os dois atores do sistema, o médico e o profissional da ARS. O utilizador acede ao sistema visualizando a informação, as várias métricas, em *dashboards*.

Caso de uso 3: Comparar informação

Apenas o profissional da ARS está envolvido neste caso de uso porque apenas ele tem acesso aos dados das várias instituições de saúde. O utilizador pode aceder ao sistema e comparar a informação que possui das várias instituições, posteriormente a ferramenta apresenta as respetivas métricas provenientes dessa comparação.

4.3 Arquiteturas propostas

Os dados terão de ser extraídos e disponibilizados ao utilizador, não conhecendo ainda a estrutura nem o tipo de dados que a base de dados de produção armazena, é possível idealizar duas formas distintas para o desenvolvimento deste sistema. A arquitetura implementada vai depender de vários fatores: o momento em que o utilizador pretende visualizar os dados, o tipo de dados que é recolhido e o tratamento que é necessário efetuar. Se o utilizador pretender visualizar a informação em tempo real, logo depois de recolhida pelo ObsCare, é necessário que a extração desses dados das fontes externas seja despoletada para iniciar todo o processamento e posterior apresentação.

No caso dos dados não necessitarem de um elevado processamento e limpeza, a ferramenta pode ser desenhada de forma a aceder diretamente às fontes externas e efetuar as transformações antes da apresentação.

A principal arquitetura apresentada contempla um DW, onde a ferramenta passaria a aceder para efetuar as análises, não tendo assim a necessidade de aceder diretamente aos dados da BD de produção.

A base de dados de produção é populada pela aplicação ObsCare, estes dados são extraídos, transformados e carregados para o DW. O DW vai disponibilizar os dados para que o Power BI possa efetuar todas as análises necessárias apresentando a informação pertinente ao utilizador.

O fluxo de dados desta arquitetura é apresentado na figura 4.2, a BD de produção e o DW são representados sob a forma de repositórios de dados, uma vez que é onde estes se encontram armazenados. Os processos identificados nesta arquitetura são a extração, transformação e carregamento dos dados para o DW e a disponibilização dos dados pelo DW para a ferramenta de análise de dados (Power BI).

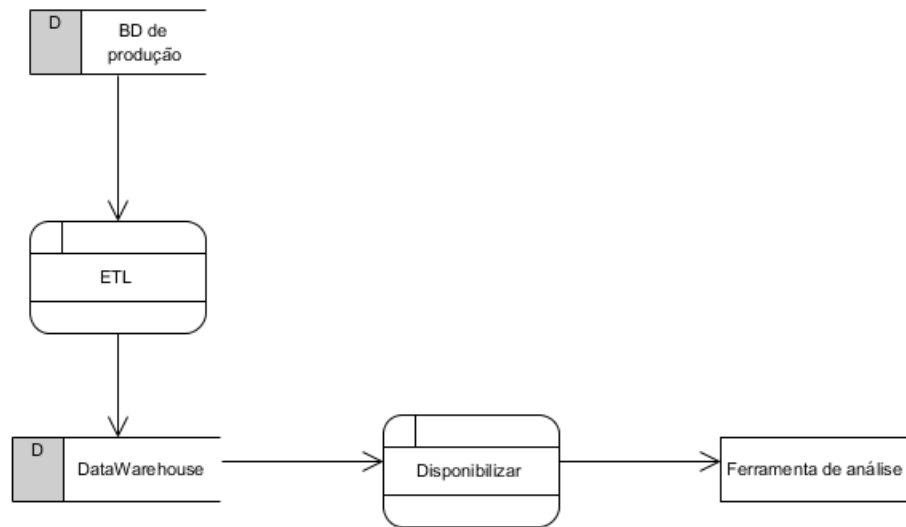


Figura 4.2: Diagrama de fluxo de dados para a arquitetura proposta.

Com esta arquitetura os dados seriam extraídos, transformados e posteriormente guardados no DW para a posterior análise. A solução proposta tem de garantir que os dados existentes no DW têm significado, são perceptíveis por qualquer utilizador, para que este sistema também possa ser utilizado e adaptado por qualquer profissional. Deste modo, é necessário garantir toda a transformação dos dados para que estes sejam compreendidos por todos os utilizadores, não só programadores.

A metodologia de Kimball vai de encontro com as necessidades anteriores, sendo esta a que se vai utilizar nesta arquitetura. Com esta metodologia vai-se garantir também que o DW não seja resistente à mudança, garantido-se assim a utilização quer de dados recentes quer de dados mais antigos, contemplando-se todo o histórico. Neste modelo proposto o DW seria constituído por um DM para o setor de obstetrícia. A arquitetura proposta contempla um DM específico para este serviço o que vai permitir que facilmente se possam adicionar novos DM de outros serviços, fazendo crescer o DW. As análises seriam assim efetuadas sobre o DM, sem nunca interferir com os sistemas fonte.

Para esta arquitetura estão contemplados quatro componentes, a BD de produção, o DW, a ferramenta de análise, e o browser. O diagrama de componentes encontra-se na figura 4.3.

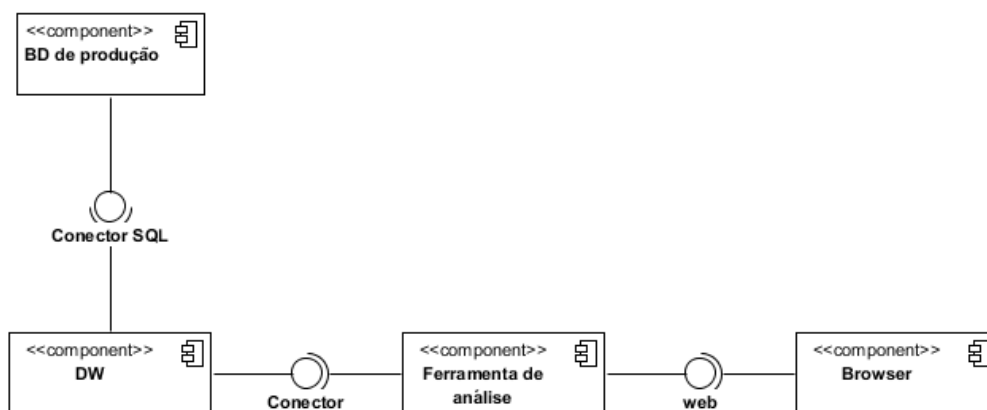


Figura 4.3: Diagrama de componentes da arquitetura proposta.

Para apresentar a forma como a arquitetura principal seria implementada a profissionais sem conhecimentos técnicos é apresentada uma vista de implementação concetual (figura 4.4). Nesta vista são apresentados os servidores necessários para o desenvolvimento deste sistema, seria um servidor para a BD de produção, outro para o DW e o outro para a ferramenta de análise, considera-se a utilização de servidores independentes para garantir a segurança dos dados.

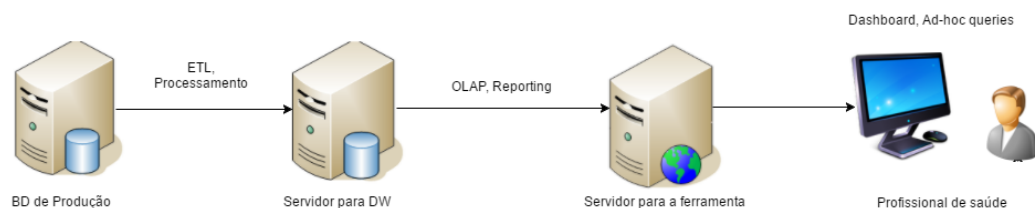


Figura 4.4: Vista concetual de implementação.

Como alternativa de arquitetura para o sistema é apresentada uma estrutura em que a ferramenta de análise efetua a conexão direta com a BD de produção. Para o desenvolvimento desta arquitetura deverá escolher-se uma ferramenta de análise que suporte BD de modelos relacionais. A BD de produção disponibiliza os dados que posteriormente seriam analisados pela ferramenta. Na figura 4.5 encontra-se representado este fluxo de dados.

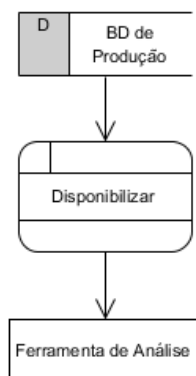


Figura 4.5: Diagrama de fluxo de dados para a alternativa de arquitetura proposta.

A base de dados de produção é populada pelo ObsCare, e nesta arquitetura a ferramenta de análise iria aceder à base de dados de produção diretamente (*queries*), sendo posteriormente a informação disponibilizada aos utilizadores nos diferentes formatos pretendidos. A ferramenta escolhida para esta arquitetura tem de respeitar alguns requisitos, por exemplo, para conseguir aceder aos dados de BD de produção, a plataforma teria de efetuar uma conexão direta com a base de dados e suportar a ligação a modelos relacionais, mesmo que isso implicasse para a ferramenta a sugestão de uma "virtualização" para um modelo dimensional, desnormalizado.

Os componentes nesta arquitetura são a BD de produção, a ferramenta de análise e o browser. A figura 4.4 apresenta o diagrama de componentes para esta arquitetura atingir as funcionalidades pretendidas.



Figura 4.6: Diagrama de componentes para a arquitetura alternativa.

4.4 Arquitetura implementada

As duas arquiteturas propostas foram avaliadas em conjunto com o cliente, concluindo-se que era necessário centralizar os dados dos diferentes hospitais, para efetuar as comparações interinstitucionais, bem como transformar os dados existentes no ambiente operacional. A centralização dos dados, a recolha de diferentes fontes leva a que seja necessário uniformizá-los para que possam ser analisados em conjunto.

Dado estas necessidades optou-se pela arquitetura que contempla um DW, composto por um DM específico do serviço de obstetrícia, deste modo, os dados serão transformados para posteriormente preencher o DW com os dados dos diferentes hospitais. Nesta estrutura, os dados vão encontrar-se centralizados onde por sua vez o Power BI irá aceder para efetuar todas as análises e gerar os relatórios necessários para os utilizadores.

A metodologia seguida, tal como indicado quando a arquitetura foi proposta ao cliente, foi a de *Kimball*. A estrutura dimensional vai permitir uma maior facilidade na interação dos utilizadores com os dados armazenados, sendo que com esta metodologia o próprio modelo é mais perceptível para todos os utilizadores. Para além disto, esta dissertação centra-se no serviço de obstetrícia, no entanto, esta abordagem permite que posteriormente, caso necessitem, possam implementar outros DM de forma independente para obtenção de outras métricas, fazendo crescer o DW.

O desenho do modelo do DW tem de ter em conta as métricas que os utilizadores necessitam de obter (algumas delas já referenciadas na tabela 2.2), no entanto, existem muito mais indicadores que se podem obter so serviço de obstetrícia.

Alguns destes indicadores podem ser referentes a:

- Produção e de qualidade dos dados, onde por exemplo se obtêm indicadores que contabilizam o número de partos e o número de recém nascidos;
- Utilização. Onde se efetua a contabilização do número de internamentos, consultas, urgências efetuadas naquele serviço;
- Partos Vaginais. Neste indicador são contabilizados, por exemplo, os partos vaginais auxiliados por outros instrumentos (fórceps, ventosas);
- Cesarianas. Neste ponto são obtidos indicadores nas diferentes técnicas utilizadas na cesariana. Em que situação é que este tipo de parto ocorreu, se por urgência ou foi uma situação programada;
- Complicações. Contabilização do número de ocorrências nas complicações que poderão existir;
- Recém nascido. Neste tópico podem ser obtidas contagens do número de recém nascido com malformações, por exemplo.

Tal como referido o ObsCare guarda informação relativa a todo o acompanhamento da utente durante a sua gravidez, deste modo são efetuados registos desde o início da gravidez até ao parto. A primeira vez que a utente se desloca ao serviço é efetuado o registo da utente, onde são guardados os seus dados pessoais, e também é criado um registo associado à gestação onde são guardados os dados relativos à gravidez (por exemplo, tipo de gravidez), bem como os profissionais de saúde que acompanham essa gestação.

Ao longo da gestação a utente irá deslocar-se ao hospital, e para cada deslocação é despoletado um evento, um episódio. Dependendo dos motivos para os quais a utente se desloca ao hospital vai originar diferentes tipos de episódio, que são: consulta; internamento; urgência. Em cada episódio são registadas diferentes informações, como por exemplo, a data da entrada e de saída do hospital, o motivo daquele episódio. Todos os episódios gerados durante a gestação da utente ficam associados à gestação, para que assim se possa efetuar toda a monitorização dessa gestação.

Ao longo de todos estes episódios é gerado um documento, que dependendo da situação é gerado um tipo de documento específico para esse evento. No caso de ser um episódio em que ocorreu o parto, é associado um documento do tipo IDP.3 que contém assim todos os dados relativos ao parto.

Quando ocorre o parto, para além de ter um documento específico, tem também um registo do recém nascido, esta informação é muito importante. Associado ao recém nascido tem-se

variadas informações, entre elas: o tipo de parto; como o recém nascido se encontrava (apresentação) no momento do parto; o peso; o sexo; se nasceu vivo ou morto.

Para efetuar a modelação do DW analisou-se todo este fluxo de dados, o ambiente operacional do ObsCare. Ao efetuar esta análise constatou-se que existiam indicadores que eram obtidos de métricas em diferentes níveis, desde a contabilização de registos por tipo de episódio (nível do episódio), contabilização de partos (nível do documento), contabilização de recém nascidos ou de determinadas características deste (nível do recém nascido).

A modelação da estrutura que vai armazenar os dados, o DW foi desenhada tendo em conta todas estas métricas, mesmo que estas não sejam todas calculadas ao longo desta dissertação, será possível obtê-las numa fase mais avançada sem necessitarem de fazer alterações ao DW.

4.4.1 Modelação do DW

Pela abordagem de Kimball, metodologia utilizada nesta dissertação, o DM é específico de uma área de negócio, e para a elaboração do modelo dimensional é necessário avaliar-se quatro pontos:

- Processo de negócio;
- Nível de granularidade;
- Definição das Dimensões;
- Identificação das medidas, factos numéricos.

A **área de negócio** identificada inicialmente tinha por base o acompanhamento do serviço de obstetrícia na obtenção de diferentes indicadores, por exemplo a contabilização do número de partos. O DM criado vai ser específico para o serviço de obstetrícia, onde cada linha da tabela de factos representará a informação dos documentos, que poderá referenciar um parto ou não.

O **nível de granularidade** da tabela de factos representa o nível ao qual se pretende obter métricas, como já referido, podem ser obtidos indicadores em diferentes níveis, sendo que o considerado na modelação foi o nível mais baixo de granularidade, as complicações da gravidez. Cada linha da tabela de factos representa a informação de um documento, que por sua vez tem as complicações da gestação associadas, sendo que quando ocorrem complicações na gestação esse número de linhas é multiplicado pelo número de complicações desse documento. Numa gestação pode ocorrer o nascimento de um bebé e ocorrerem várias complicações, a tabela de factos vai conter informação até às complicações, para obtenção das diferentes complicações ocorridas nos partos.

Quando ocorrer o parto será gerado um tipo de documento específico, diferente de todos os outros e será guardada informação desse parto, e do recém nascido, bem como das diferentes complicações ocorridas.

Este nível de granularidade é importante para os profissionais porque vai permitir a obtenção de indicadores associados às complicações da gravidez, aos partos, recém nascidos, e numa fase posterior se assim o entendam poderão obter outros indicadores de níveis de granularidade mais altos (episódio, gestação).

A baixa granularidade da tabela de factos, permite ao utilizador maior capacidade analítica, uma vez que também permite a obtenção de indicadores nos níveis superiores.

Para **definir as dimensões**, tal como já foi referido, foi efetuado um estudo ao ambiente operacional, bem como o levantamento das análises que deverão ser feitas neste serviço. Com base nessas análises, e de forma a dar resposta às necessidades dos utilizadores foram encontradas as seguintes dimensões:

- DimData;
- DimLocal;
- DimTipoDocumento;
- DimTipoParto;
- DimMotivos;
- DimTipoEpisódio;
- DimComplicacoes;
- DimTipoGravidez;
- DimApresentacao;
- DimDoentes.

DimData

A tabela DimData representa a dimensão da data. É importante que os dados armazenados no DW possam ser situados no tempo, quer para verificação de histórico, quer pela própria necessidade de pesquisa de dados do serviço em questão. Este fator é importante para se efetuarem as análises e as respetivas comparações. Na criação desta dimensão definiu-se que o nível de detalhe mais baixo seria o dia, sendo que a chave primária desta tabela não é uma sequência mas o código representativo do dia, para facilitar no momento da pesquisa. Por exemplo, o registo nesta dimensão do dia 1 de janeiro de 2017, ficava com a chave de 20170101, sendo facilmente perceptível, pelo código, o dia a que corresponde o registo.

Esta dimensão ao contrário de todas as outras não vai ter apenas uma chave na tabela de factos, mas sim duas. Estas duas chaves vão representar a data de inicio do episódio e data de nascimento do recém nascido, são as duas chaves da dimensão uma vez que pode ser necessário efetuar análises sobre elas de forma independente.

Na tabela de fatos as colunas têm o nome de SK_DATA_INICIO_EPISODIO e SK_DATA_NASCIMENTO_RN, colunas essas que efetuam a ligação com a SK_DATA, este mapeamento encontra-se representada na figura 4.7.



Figura 4.7: Relação entre a dimensão da data e a tabela de factos.

DimLocal

Cada registo está associado a um local, um hospital, é importante ter uma tabela com este mapeamento de forma a que facilmente se possa obter mais informações sobre o local em questão. A tabela DimLocal vai assim conter este mapeamento, esta informação é útil nas análises efetuadas, por exemplo, para se obter o número de partos efetuados num determinado hospital, ou mesmo comparar esta métrica entre os diferentes locais.

Tal como já foi referido o DW vai centralizar a informação do serviço de obstetria dos diferentes hospitais, uma vez que todos estes registos se encontram na mesma estrutura é fundamental conseguir identificar de forma clara a que instituição pertencem.

Esta dimensão foi efetuar a ligação com a tabela de factos por uma única chave SK_LOCAL, tal como é representado na figura 4.8

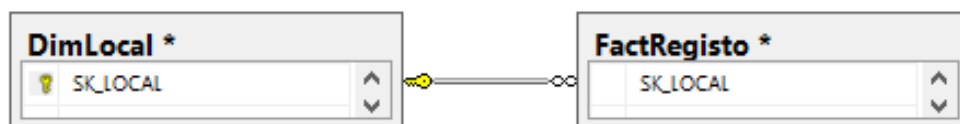


Figura 4.8: Relação entre a dimensão do local e a tabela de factos.

DimTipoDocumento

Em todos os episódios é preenchido um documento, no entanto, os documentos são caracterizados pelo seu tipo. Dependendo do episódio é gerado um documento específico, por exemplo, no episódio em que ocorre o parto é produzido um documento do tipo IDP.3. Esta característica do documento vai assim ajudar na caracterização do registo, podendo-se efetuar análises de dados por esta característica.

A tabela DimTipoDocumento vai conter o mapeamento destes códigos com as respetivas descrições podendo-se, numa fase mais avançada, utilizar-se esta dimensão nas análises. Esta dimensão vai ligar com a tabela de factos através da coluna SK_Tipo_Documento, este mapeamento encontra-se representado na figura 4.9.



Figura 4.9: Relação entre a dimensão dos tipos de documento e a tabela de factos.

DimTipoParto

Outra perspetiva de análise dos dados é pelo tipo de parto, a tabela DimTipoParto vai ter a descrição dos diferentes tipos de parto e a respetiva codificação. Esta dimensão vai permitir filtrar os dados e análises consoante o tipo de parto ocorrido, com esta dimensão será possível, por exemplo, contabilizar o número de cesarianas (tipo de parto presente nesta dimensão). O mapeamento com a tabela de factos é efetuado pela chave SK_TipoParto, a tabela de factos vai conter apenas uma chave desta dimensão tal como é representado na figura 4.10.

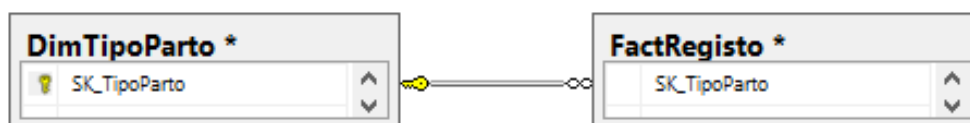


Figura 4.10: Relação entre a dimensão dos tipos de parto e a tabela de factos.

DimMotivos

Um dos indicadores mais importantes é o número de cesarianas, mas para se efetuar uma cesariana tem de se associar sempre um motivo. Estes motivos por sua vez também são alvos de análise e monitorização. A tabela DimMotivos vai efetuar o mapeamento dos códigos dos motivos com as respetivas descrições. Este mapeamento é de extrema relevância quer para a justificação da execução desse parto como também para obtenção de indicadores específicos dos motivos, por exemplo, percentagem de cesarianas efetuadas devido a cirurgia uterina prévia. Esta dimensão vai estar ligada com a tabela de factos pela chave SK_Motivo, como se encontra representado na figura 4.11.

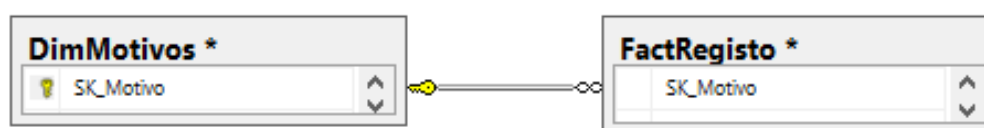


Figura 4.11: Relação entre a dimensão dos motivos e a tabela de factos.

DimTipoEpisódio

Outra dimensão é a DimTipoEpisódio, tal como já foi referido, os episódios podem ser de três tipos: urgência; consulta; internamento. O serviço de obstetrícia necessita de analisar estes três tipos, agrupando os dados e obtendo indicadores sobre cada tipo de episódio. A dimensão vai permitir assim analisar os dados consoante os diferentes episódios.

A tabela de factos vai apenas conter uma chave desta dimensão, a SK_TipoEpisodio, este mapeamento encontra-se representado na figura 4.12.

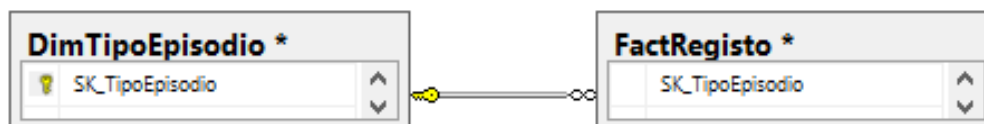


Figura 4.12: Relação entre a dimensão dos tipos de episódio e a tabela de factos.

DimComplicacoes

O parto efetuado pela forma de cesariana não deixa de ser uma intervenção cirúrgica e como tal tem consequências, complicações. Para se efetuar o mapeamento destas complicações foi criada a dimensão DimComplicacoes, o que vai permitir analisar quais as complicações mais recorrentes desta cirurgia, bem como a obtenção de outros indicadores.

A tabela de factos tem a chave SK_Complicacoes que liga com a chave da dimensão, esta conexão encontra-se representada na figura 4.13.



Figura 4.13: Relação entre a dimensão das complicações e a tabela de factos.

DimTipoGravidez

A caracterização dos registos também se reflete no tipo de gravidez da utente, fator importante na análise. A tabela DimTipoGravidez contém o mapeamento dos tipos de gravidez com as respetivas descrições, este mapeamento é efetuado com a presença da chave SK_TipoGravidez na tabela de factos que liga com esta dimensão, tal como é demonstrado na figura 4.14.

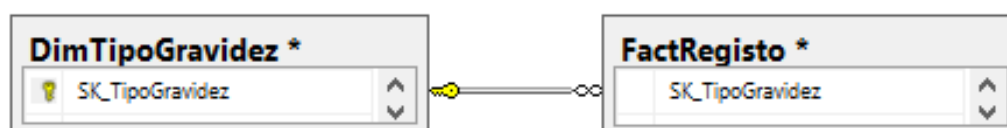


Figura 4.14: Relação entre a dimensão dos tipos de gravidez e a tabela de factos.

DimApresentacao

No momento do parto é guardado como o recém nascido se apresenta, a posição em que se encontra. O mapeamento entre o código da apresentação e a respetiva descrição encontra-se guardado nesta dimensão, DimApresentacao. Com este mapeamento é possível perceber como essa apresentação pode ter influenciado o tipo de parto. Com esta tabela o utilizador pode efetuar análises e caracterizar os partos pelos atributos desta dimensão. A tabela de factos vai conter a chave SK_Apresentacao que efetua a ligação com esta dimensão, a figura 4.15 apresenta este mapeamento.



Figura 4.15: Relação entre a dimensão da apresentação e a tabela de factos.

DimDoentes

A tabela DimDoentes vai guardar a informação das utentes de cada hospital, nesta dimensão encontra-se uma coluna que indica de que hospital pertence a utente. Com estes dados pode-se acompanhar toda a gestação dessa utente bem como efetuar análises específicas com os atributos desta dimensão, por exemplo, contabilização do número de cesarianas por idade, para verificar em que faixa etária este tipo de parto é mais predominante. A tabela de factos vai ter assim a chave SK_Doente que representa a chave desta dimensão, este mapeamento encontra-se representado na figura 4.16.



Figura 4.16: Relação entre a dimensão dos doentes e a tabela de factos.

O ultimo passo, como refere Kimball, é a **identificação dos factos numéricos** que estão presentes na tabela de factos. As variáveis numéricas consideradas para se efetuarem as análises são: o peso da utente anterior à gestação, altura da utente, idade gestacional do bebé no momento do episódio, peso do recém nascido no momento do parto, semanas de gestação do recém nascido na altura do parto, perímetro cefálico do recém nascido.

Para além dos valores numéricos são também contempladas flags que vão permitir analisar e efetuar contagens sobre os registos que são:

- flag_gesta_aborto. Esta flag vai indicar se aquela gestação originou um aborto;
- flag_rn_vivo. Indica se o recém nascido nasceu vivo ou não;
- rn_sexo. Sigla que representa o sexo do recém nascido;
- flag_rn_malform. Esta flag indica se o recém nascido nasceu, ou não, com mal formações;

- `flag_inicio_trabalho parto`. Flag que indica o trabalho de parto da utente, se foi induzido ou espontâneo.

Para a definição deste modelo foi utilizado o esquema em estrela, em que todas as dimensões se relacionam com a tabela de factos. Este modelo oferece melhor desempenho nas consultas efetuadas, uma vez que a necessidade de efetuar joins é mais reduzida e também é um modelo mais perceptível para o utilizador final.

Cada linha da tabela de factos, `FactRegisto`, representa um documento (com as respetivas complicações), contendo as diferentes métricas e as diferentes características necessárias para a análise dos dados. Esta tabela de factos relaciona-se com todas as dimensões tendo assim todas as chaves estrangeiras referentes a cada dimensão. A relação com as dimensões permite a caracterização das métricas pelos atributos presentes nas dimensões. A modelação deste DM específico para o serviço de obstetrícia encontra-se representado na figura 4.17 . A tabela de factos encontra-se ao centro do esquema conectada com todas as dimensões, sendo que cada chave da dimensão pode ocorrer várias vezes na tabela de factos tratando-se assim de uma relação de 1 para n. Tal como referido, existe uma dupla relação com a dimensão da data, relação essa que se encontra também representada na figura.

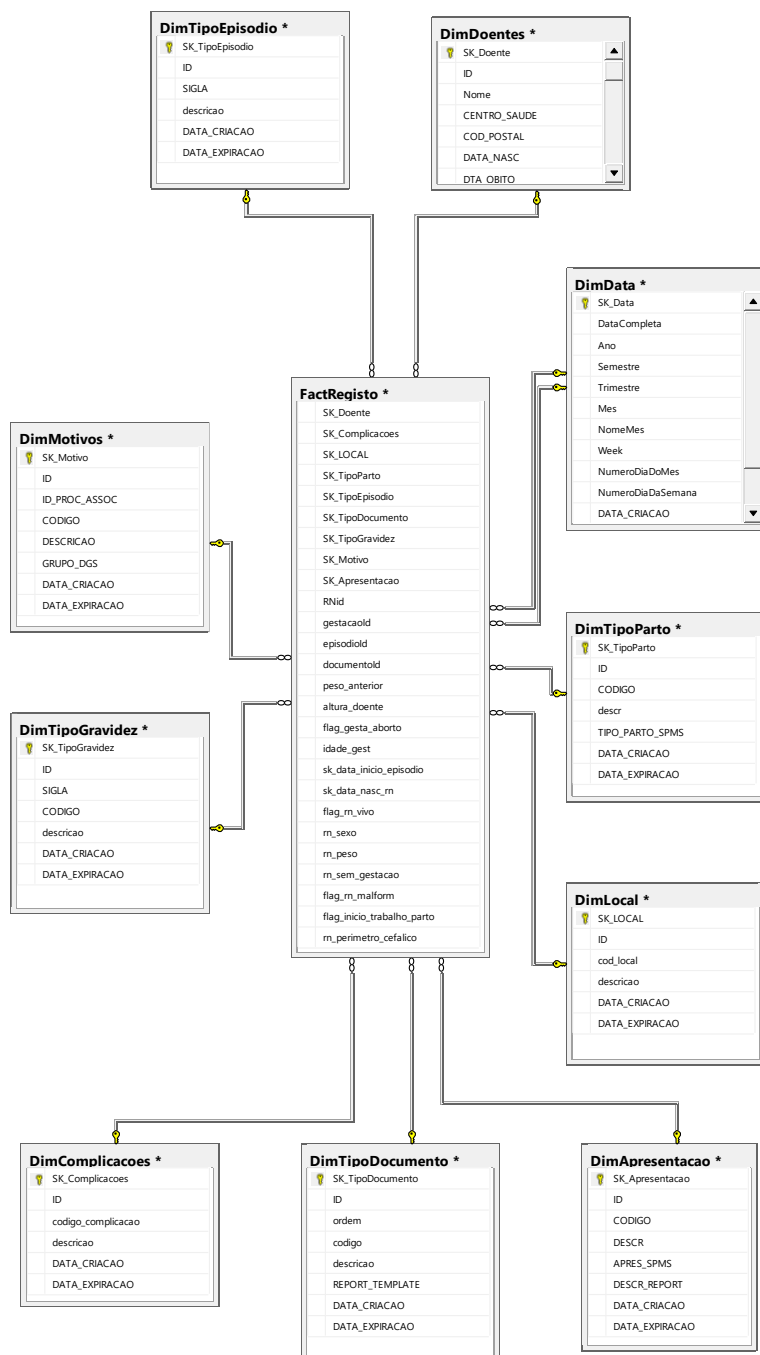


Figura 4.17: Modelo do DM implementado para o serviço de obstetrícia

4.4.2 Implementação do DW

O carregamento do DW é efetuado por partes e de forma sequencial. Para preencher a tabela de factos as dimensões têm de apresentar os dados mais atualizados, e para isso estas tabelas são carregadas em primeiro lugar, seguindo-se o carregamento da tabela de factos.

Tal como já referido, os dados provêm de diferentes hospitais e para agilizar todo este processo de ETL e carregamento do DW existe uma tabela de configurações, CONFIG_DW. Nesta tabela é indicado as parametrizações necessárias para cada hospital, nomeadamente, o período em que os dados poderão ser extraídos dos hospitais.

Será possível parametrizar o intervalo de tempo (por exemplo, se a extração será efetuada todos os dias (um dia de intervalo), ou uma vez por semana(sete dias de intervalo) e a hora dessa extração, hora em que se acederá ao hospital. Esta parametrização é de extrema importância para os hospitais uma vez que assim é possível agendar este processo em conformidade com os vários processos internos do hospital.

As dimensões (exceto a dimensão dos doentes) são tabelas com codificações iguais em todos os hospitais, e como tal, terão uma parametrização específica, onde é indicado o hospital de referência para se efetuar a extração dos dados.

Para a parametrização do carregamento da dimensão da data, dimensão estática, não é contemplado nenhum hospital como fonte de dados. No entanto, na tabela de configurações é indicado o período de carregamento, sendo que nesta fase para o carregamento do DW é apresentado um carregamento mensal, como forma de preencher a tabela com os códigos para os meses enviados por parâmetro.

O carregamento da dimensão dos doentes e a tabela de factos será efetuado com base nos dados de todos os hospitais, definindo-se também o período e a hora da extração dos dados na tabela de configurações.

Carregamento das dimensões

Os dados guardados nas dimensões podem sofrer alterações de uma forma imprevisível e como tal é necessário perceber quais os atributos que realmente têm relevância para se guardar histórico, bem como perceber-se como esse armazenamento de histórico é efetuado. Para guardar estas transformações nas tabelas de dimensões existem três formas mais vulgares: tipo 1; tipo 2; tipo 3.

Quando um atributo do tipo 1 sofre alterações, o registo anterior é atualizado ficando assim o atributo com o valor atualizado. Desta forma não é guardado histórico uma vez que não é possível obter-se o valor que se encontrava anteriormente à atualização.

Um atributo do tipo 2, indica que quando sofre alteração é adicionada uma nova linha. Quando é verificado que esse atributo sofreu alterações, o registo anterior é fechado e é inserido um novo registo com o valor desse atributo atualizado. Com este tipo é possível fazer uma reconstrução do histórico, podendo obter-se o valor daquele atributo num determinado período temporal.

Para os atributos do tipo 3, as alterações são guardadas sob a forma de uma nova coluna na tabela. Quando o atributo sofre alterações é adicionada uma nova coluna com o valor

atualizado, tendo-se assim as duas colunas, uma com o valor antigo e outra com o valor mais recente.

Os diferentes tipos foram avaliados e para o carregamento das dimensões deste DW foram considerados apenas os atributos do tipo 1 e do tipo 2.

Para o carregamento das dimensões que utilizam apenas um hospital como fonte de dados (DimTipoGravidez, DimTipoEpisodio, DimTipoDocumento, DimMotivos, DimApresentacao, DimTipoParto, DimLocal, DimComplicacoes), foi criado um procedimento independente onde é passado por parâmetro a ligação do hospital parametrizado para aquela extração. O volume de dados nestas dimensões é reduzido, tanto em número de registos, como em número de atributos, como tal, definiu-se que para estas dimensões se iria guardar histórico de todos os atributos, em que todos eles seriam do tipo 2.

Nestas dimensões qualquer que seja o atributo alterado, o registo anterior será fechado preenchendo-se a data de expiração e criado um registo novo com o valor do atributo atualizado.

A dimensão da data é estática, os registos desta tabela não sofrem alterações. No carregamento inicial, em que a dimensão ainda não tem nenhum registo, é verificado qual o valor mínimo da data de registo dos episódios no sistema operacional (é utilizado o hospital de referência), com essa data é preenchida esta dimensão até ao mês seguinte, com todas as variáveis correspondentes aos vários dias deste período temporal. Depois deste carregamento inicial ocorrer, o processo correrá mensalmente onde preencherá, com os dados do mês enviado por parâmetro, os diferentes atributos constituintes da dimensão.

Para o mês enviado como parâmetro é obtido o semestre, o trimestre, são calculados todos os dias desse mês, verificando-se qual o dia da semana a que cada um corresponde e qual o número da semana do ano esses dias se enquadram. Depois de se obter todas estas informações relativas ao mês, os dados são inseridos na dimensão, sendo possível, posteriormente, analisar os factos ao nível de granularidade do dia.

Para efetuar o carregamento da DimDoentes, analisou-se para cada atributo a necessidade, ou não, de armazenar histórico. Como é uma dimensão maior, que representa todos os utentes dos vários hospitais houve a necessidade de se analisar cada atributo para identificar o tipo de registos. Após análise verificou-se que existem atributos que não têm necessidade de armazenar histórico, podendo-se efetuar o update direto no campo (tipo 1) que não iria ocorrer perda de informação relevante para as análises. Os atributos identificados como do tipo 1 são:

- Nome;
- Des_Postal;
- Data_Nasc;
- Estado_Civil;
- Mae;
- Pai;
- Profissao;
- Seq_Postal;

- SEXO;
- Centro_Saude_Outro;
- Pais_Residencia;
- Data_Obito.

Nesta análise foram encontrados também os atributos que era necessário armazenar histórico, campos esses que poderão ser relevantes nas análises efetuadas aos dados. Para estes atributos foi definido o tipo 2, que são:

- Nacionalidade;
- Centro_Saude;
- Cod_Postal;
- Morada;
- Num_Benef;
- Num_Utente;
- Subsistema;
- Medico_Centro_Saude;
- Cod_Subsistema;
- Num_Processo_Orig.

Depois de efetuada a análise de cada atributo procedeu-se à criação de um processo que efetua o carregamento das utentes. Neste processo é verificado se ocorreram alterações a registos existente no DW ou se são registos novos. No caso de ocorrerem alterações a utentes já presentes no DW procedesse ao processamento desses registos consoante o atributo que sofreu alterações, atributo do tipo 1 atualiza-se diretamente o registo, atributo do tipo 2 fechasse o registo anterior e insere-se um novo.

Carregamento da tabela de factos

O serviço de obstetrícia efetua o acompanhamento das utentes durante a gestação, durante o parto e após o parto. Para haver este acompanhamento é efetuado o registo da gestação da utente onde é efetuado, por exemplo, o registo do tipo de gravidez, características da mãe (se tem o plano de vacinação cumprido, o peso anterior à gestação, grupo sanguíneo).

Uma gestação tem uma duração de cerca de nove meses, durante este período a gestante ¹ necessita de ser acompanhada, tendo de efetuar deslocações ao local onde está a ser acompanhada (consultas), no entanto, caso a utente apresente algum problema poderá ter de se deslocar não por questões programadas, mas sim pelas suas condições de saúde. De cada vez que a utente se desloca ao serviço é criado um evento, um episódio, que dependendo das condições da deslocação o tipo de episódio assume diferentes valores (consulta, internamento, urgência). Deste modo, uma gestação tem vários episódios em diferentes períodos temporais, ficando associado a estes diferentes atributos nomeadamente a data do início do episódio.

¹Pessoa que está em gestação.

Para cada episódio criado é associado um documento, este documento é preenchido pelo profissional de saúde com informação mais detalhada daquele episódio e daquela gestação, por exemplo fica associado ao documento a idade gestacional no momento daquele episódio. O documento criado é caracterizado também pelo seu tipo, uma vez que dependendo da situação é preenchido o respetivo documento, por exemplo, se ocorrer o parto naquele episódio é preenchido um documento do tipo dados do parto ('IDP.3'), caso ocorra uma cesariana é preenchido um documento referente ao registo cirúrgico (realizar uma cesariana é equivalente a efetuar uma cirurgia). quando ocorre uma cesariana há um conjunto de complicações que podem advir da cirurgia a que a utente está sujeita, por cada cesariana pode existir mais que uma complicação, no entanto todas estas complicações ficam associadas ao mesmo documento (documento esse representante do registo cirúrgico), as complicações ficam também associadas à gestação.

O processo de carregamento da tabela de factos é efetuado de forma faseado, inicialmente são extraídas as complicações das várias gestações (complicações associadas aos documentos), e de seguida são extraídos os dados das restantes tabelas principais (gestação, episódio, documento, doente).

Os dados são extraídos para o DW e posteriormente são transformados e relacionados com as dimensões, de forma a efetuar a inserção dos dados na tabela de factos.

4.4.3 Configuração da visualização dos dados

O Power BI permite a conexão com várias fontes de dados, para este caso de estudo, a fonte de dados é uma base de dados oracle, e para que a ferramenta consiga efetuar a ligação com esta fonte é necessário a instalação de um conjunto de bibliotecas que permitam o acesso a instâncias de bases de dados oracle.

Foi instalado um *oracle instant client* com a versão em concordância com a versão do Power BI, isto é, se a ferramenta for instalada para 64 bits o *instant client* também tem de ser para 64 bits.

Depois desta configuração a conexão da ferramenta com a base de dados é efetuada de forma intuitiva e simples. Para se efetuar a visualização dos dados deve-se selecionar a fonte de dados, neste caso, base de dados oracle, como é apresentado na figura 4.18.

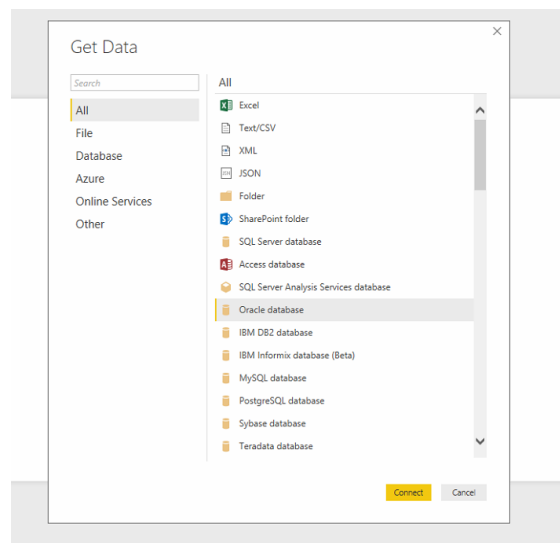


Figura 4.18: Escolha da fonte que possui os dados, base de dados oracle.

A conexão é efetuada para a máquina configurada no *instant client*, caso esta possua user-name e password é necessário indicar na conexão do Power BI 4.19.

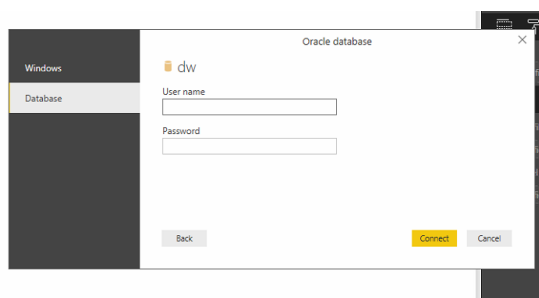


Figura 4.19: Autenticação da base de dados.

A ferramenta depois de efetuar a conexão com a base de dados apresenta todas as tabelas disponíveis para que o utilizador possa escolher às quais pretende aceder, neste caso apresenta as várias dimensões e a tabela de factos, tal como é apresentado na figura 4.20.

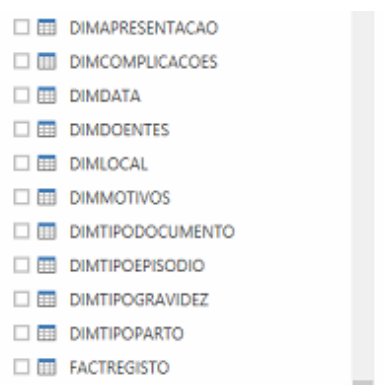


Figura 4.20: Tabelas disponíveis para visualização no Power BI.

No modelo de dados apresentado para o DW, figura 4.17, é apresentada apenas uma dimensão para a data, no entanto, na tabela de factos tem-se duas chaves estrangeiras para a data, o facto de ter duas chaves não quer dizer que se tenham de ter duas dimensões exatamente iguais do lado da base de dados para a data. Desta forma, no Power BI foi carregada a dimensão da data para cada uma das chaves da tabela de factos. As dimensões ficaram assim associadas às respetivas chaves SK_DATA_INICIO_EPISODIO e SK_DATA_NASCIMENTO_RN.

O Power BI permite a criação de relações entre as diferentes tabelas, permitindo assim obter-se o mesmo modelo, em estrela, do lado da ferramenta. Esta estrutura é fundamental para a ferramenta conseguir efetuar todos os cálculos e navegações pelas tabelas (filtragens cruzadas), pois se a estrutura das relações efetuar algum tipo de *loop*, isto é, todas as tabelas estarem relacionadas entre si, vai originar ambiguidade de dados, não fornecendo valores viáveis nas filtragens cruzadas (Microsoft, 2017a).

Para além da possibilidade de criação do modelo, a ferramenta permite a sua exibição, onde mostra todas as relações, a direção dessa relação e a cardinalidade (Microsoft, 2017a), para o caso em estudo, o modelo obtido encontra-se representado na figura 4.21, o modelo obtido vai de encontro ao modelo definido para o DW, modelo em estrela.

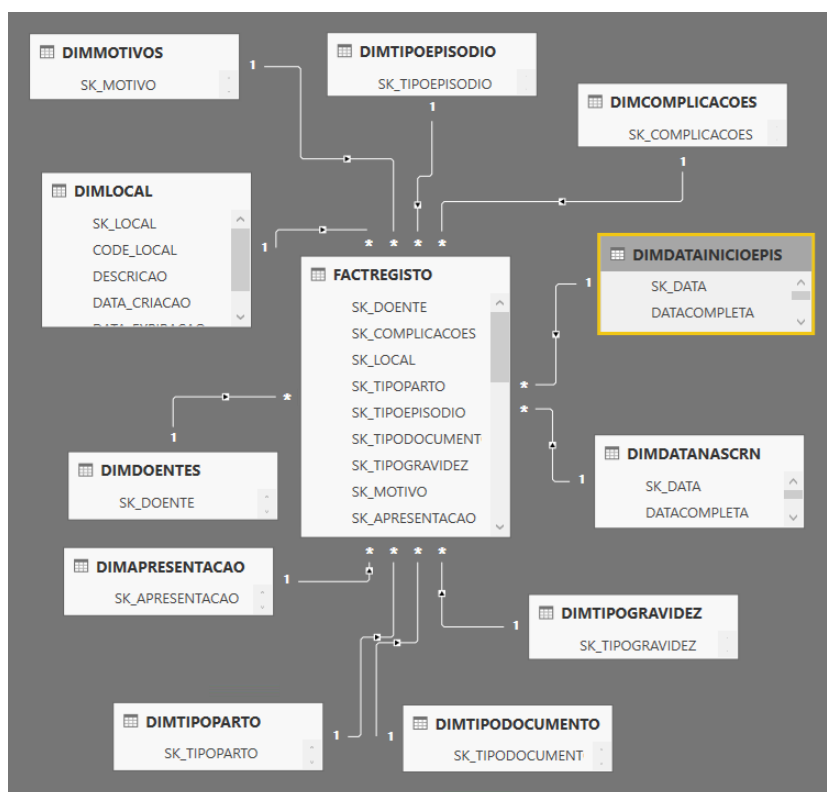


Figura 4.21: Relações entre as tabelas representadas no Power BI.

Ao contrário de outras aplicações concorrentes de análise de dados para esta não é necessária a construção de um cubo OLAP. Como a ferramenta permite a criação das relações, pode-se assim filtrar os dados com base em todas as dimensões, permitindo também a construção de hierarquias do lado da ferramenta, não sendo necessária a construção do cubo. Por exemplo, se for criada uma hierarquia do lado do Power BI na dimensão da data, com o

ano, semestre, trimestre, mês, análises que utilizem esta hierarquia poderão ser efetuadas com um *drill down* até ao mês.

Tal como já foi apresentado nas características da aplicação, esta possui elevadas capacidades de processamento, fornecendo um conjunto de expressões e fórmulas que permitem uma análise muito mais simplificada e intuitiva dos dados. O Power BI permite adicionar novas colunas às respetivas tabelas e novas métricas não alterando o modelo de dados, estes campos adicionados ficam disponíveis na ferramenta, em memória, para posterior utilização nas diferentes análises e visualizações.

Para partilhar os relatórios com outros utilizadores é necessário publicá-los no Power BI Service, no entanto, para partilhar o conteúdo com outros utilizadores o proprietário do relatório necessita de uma licença Pro do Power BI. Os dados visualizados pelos outros utilizadores são os mesmos aos quais o utilizador que criou o relatório tem acesso, a menos que este aplique um nível de segurança sobre o conjunto de dados.

Outra funcionalidade da partilha dos dados, controlada pelo proprietário do relatório, é permitir que os utilizadores com os quais partilhou os relatórios/*dashboards* possam também partilhar com outros utilizadores. Uma das formas de partilhar o relatório é pela aplicação, mas os outros utilizadores com os quais se vai partilhar também têm de possuir a licença Pro. Dependendo da forma como é partilhado, pode ser partilhado num grupo com vários elementos da mesma organização ou criando uma aplicação (conjunto de relatórios e *dashboards*) onde todos podem contribuir para o desenvolvimento de novos relatórios. Para a partilha individual de um determinado relatório, os utilizadores com os quais são partilhados os relatórios apenas podem visualizar os dados e interagir com eles, não podem editar os relatórios (Microsoft, 2017b).

Pode-se também partilhar os relatórios com recurso ao *sharepoint*, sendo possível ao proprietário do relatório incorporá-lo numa página do *sharepoint*, onde os utilizadores autorizados poderão ver e interagir com os relatórios. A figura 4.22 apresenta um dos relatórios criado para o caso de estudo partilhado com recurso ao *sharepoint*, nesta plataforma o utilizador pode também interagir com os dados e filtrá-los com recurso a filtros que não existam no *layout* do relatório.

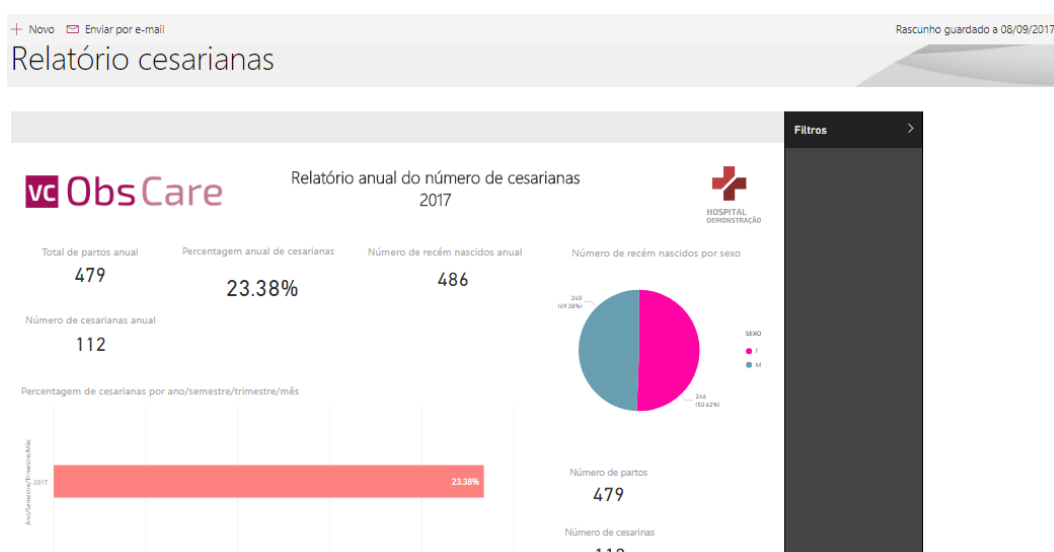


Figura 4.22: Relatório principal de taxa de cesarianas partilhado no *share-point*.

É também possível partilhar os relatórios através da publicação do relatório na *web*, sendo gerado um *Uniform Resource Locator* (URL) que pode ser incorporado em sites públicos e partilhado com os vários utilizadores. Em todas estas formas de partilha o relatório partilhado fica sincronizado com o relatório de origem.

Em todas as formas de partilha tem de se ter especial atenção à segurança dos dados, e à forma como estes ficam acessíveis para análise do lado do Power BI. Para que os dados não migrem para a *cloud* no momento da publicação, tem de se usar uma conexão às fontes por *direct query* isto é, no momento da visualização dos dados a ferramenta conecta com as fontes e mostra a informação, não ficando assim guardada em nenhum local (Microsoft, 2017c).

4.5 Conclusão

Depois de identificado o problema, foi elaborada uma solução que permite dar-lhe resposta. Como tal, foi implementado um DW para o serviço de obstetrícia.

A implementação do DW permitiu a centralização dos dados referentes a diferentes hospitais, sendo possível também armazenar-se histórico. A modelação foi efetuada segundo um modelo dimensional, para que todas as análises sejam efetuadas de forma rápida e mais perceptível para o utilizador. Toda a parte de modelação dos dados e o ETL foi efetuado numa base de dados oracle disponibilizada para o efeito.

Com esta centralização dos dados, a ferramenta de análise acede ao DW tendo assim todos os dados necessários para as análises e as comparações dos dados. Através desta centralização da informação é possível disponibilizar os vários indicadores e as diferentes análises.

Capítulo 5

Análise dos resultados

Neste capítulo são apresentados os vários indicadores obtidos com a solução desenvolvida. Estes indicadores foram definidos ao longo da dissertação com base nas necessidades do cliente e pesquisa bibliográfica.

Nesta secção é também efetuado o mapeamento dos resultados obtidos com os requisitos definidos para a solução.

5.1 Indicadores

A taxa de cesarianas é o principal indicador internacional de qualidade do serviço de obstetrícia (WHO, UNFPA e UNICEF, 2009; Long et al., 2015) e como tal é imperativo que esta seja baixa de forma a garantir uma maior qualidade. A cesariana é uma cirurgia, desta forma, tem complicações associadas, para além disto, podem ocorrer outras complicações para a mãe e para o bebé. A diminuição desta taxa vai diminuir diretamente estas complicações havendo assim um contributo para a melhoria da qualidade dos cuidados de saúde prestados neste serviço.

Para compreender porque motivo a taxa de cesarianas é elevada e de que forma é possível reduzi-la, é necessário compreender os principais motivos pelos quais as cesarianas acontecem e qual a sua evolução.

Recorrendo a uma ferramenta de análise de dados, neste caso o Power BI, é possível verificar a evolução dos principais indicadores que afetam a taxa de cesarianas facilitando aos profissionais de saúde a compreensão do porquê da taxa, a sua tendência e quais as medidas que poderão implementar para reduzir a mesma.

5.1.1 Número de cesarianas anual

Uma das preocupações num serviço de obstetrícia, tal como tem vindo a ser abordado, é o número de cesarianas realizadas, e que impacto tem esse número na totalidade dos partos. No seguimento desta principal necessidade, foi definido um relatório principal em conjunto com o cliente. Este relatório apresenta métricas do ano corrente, no caso 2017, para o diretor de serviço de um hospital. Neste relatório consta o número de partos realizados nesse ano, e quais deles foram cesarianas.

A relação entre o número de partos e o número de recém nascidos não é igual, de um parto pode nascer mais que um bebê. Assim sendo, neste relatório principal deve ser apresentado também o número de recém nascidos.

De salientar que o DW possui a informação dos vários hospitais, desta forma todas as métricas têm de ser calculadas sempre por local, para se disponibilizar os dados corretos do serviço de obstetrícia a que faz parte o utilizador. Para dar resposta a este relatório foram calculadas quatro métricas do lado do Power BI:

- Número total de partos, ocorridos no ano corrente;
- Número de cesarianas, ocorridas no ano corrente;
- Percentagem de cesarianas para o mesmo ano;
- Número de recém nascidos, nascidos no mesmo ano.

Para dar resposta a esta necessidade do cliente foi criado um relatório, com base em apenas um local, Hospital Demo, utilizando os logotipos do respetivo hospital e do ObsCare. Para este relatório não foram disponibilizados filtros adicionais o que quer dizer que o utilizador, apenas tem disponível a informação presente no *layout* do relatório.

O relatório final disponibilizado ao cliente encontra-se na figura 5.1.

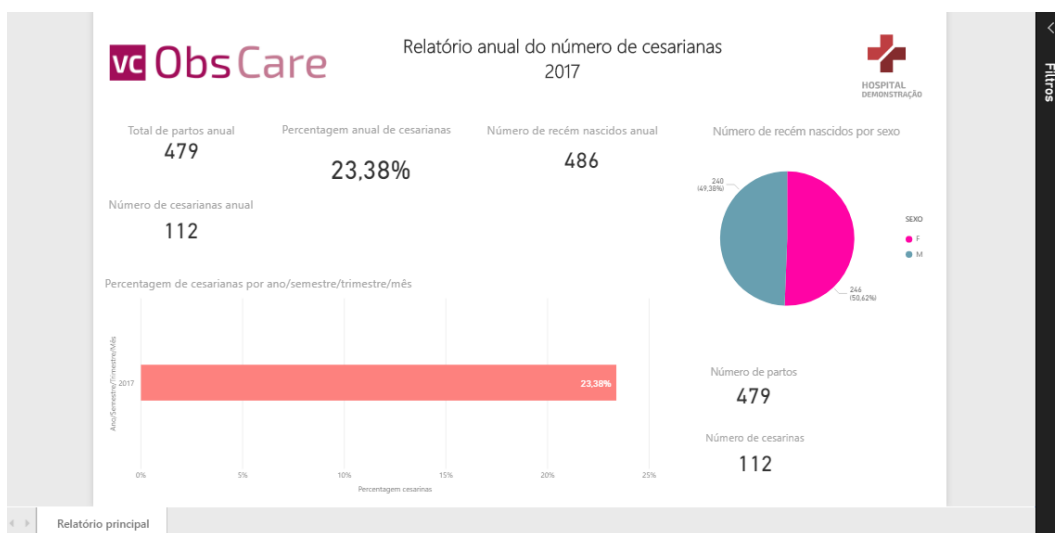


Figura 5.1: Relatório com taxas anuais fornecido ao utilizador.

Todos os valores apresentados dizem respeito ao ano corrente, para este hospital de demonstração o valor total de partos é 479, sendo que 112 foram cesarianas, apresentando assim uma taxa de cesarianas de 23,38%. De um parto pode nascer mais que um recém nascido, caso de gémeos, e desta forma é relevante saber qual foi o número de recém nascidos para o número de partos que se está a analisar, tal como esperado este valor é superior ao número total de partos. Caso o utilizador pretenda a distinção desse número de recém nascidos por sexo, poderá visualizá-lo no relatório, do lado direito deste é apresentado um gráfico circular onde é diferenciado o número de recém nascidos por género (feminino, masculino).

O utilizador ao analisar as métricas disponibilizadas pelo relatório vai conseguir verificar se o valor da taxa de cesarianas se encontra alto ou baixo. Caso apresente um valor alto o utilizador pode averiguar os motivos e tomar medidas para que esse valor baixe, sendo que

depois da implementação dessas medidas vai poder continuar a monitorizar essa taxa neste relatório.

No entanto, se o utilizador pretender analisar períodos mais específicos, por exemplo, o primeiro semestre do ano, ou o primeiro trimestre, ou comparar o primeiro trimestre do ano com o segundo para verificar a evolução ao longo do ano pode efetuar pela análise do *dashboard* apresentado à esquerda na parte inferior.

Este *dashboard* disponibiliza a taxa de cesarianas nos diferentes períodos, semestre, trimestre, mês, o utilizador tem de descer até ao nível que pretende analisar. Para a visualização dos dados em todos estes níveis foi necessário criar um hierarquia composta pelo ano-semester-trimestre-mês, com esta hierarquia é possível efetuar um *drill down* nos dados navegando nos diferentes níveis tal como se encontra apresentado nas figuras 5.2, 5.3, 5.4.

Os dados utilizados são apenas referentes ao primeiro semestre do ano, e por isso, é apenas apresentada a taxa de cesarianas para esse semestre, 23,38% , figura 5.2.

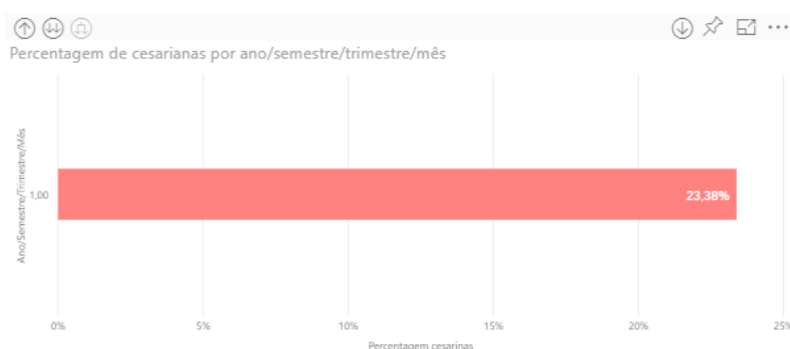


Figura 5.2: Percentagem de cesarianas para o primeiro semestre do ano corrente.

A figura 5.3, apresenta um *drill down* da taxa de cesariana ao nível do trimestre. Para o primeiro trimestre do ano é apresentada uma taxa de 23,11%. Neste período foram realizados 476 partos, em que 110 foram cesarianas, o utilizador ao clicar no primeiro trimestre consegue visualizar estes valores ao lado do *dashboard*, esses valores são assim alterados consoante o período que o utilizador seleciona.

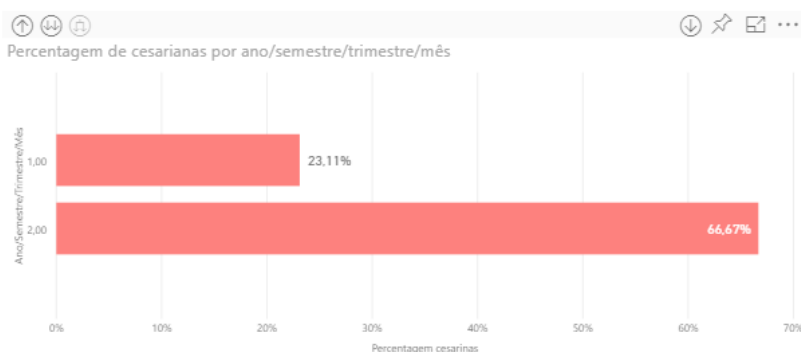


Figura 5.3: Taxa de cesarianas por trimestre.

No segundo trimestre a taxa de cesarianas apresentava um valor de 66,67%, visualizado ao pormenor esse trimestre é indicado que foram realizados três partos, em que dois deles foram cesarianas.

Uma forma de perceber como essa taxa está distribuída pelos meses do ano, é descendo ao nível do mês, figura 5.4.

Verifica-se que o mês de Maio tem uma taxa de cesarianas de 100% (este mês faz parte do segundo trimestre do ano). Com a informação fornecida no nível anterior constata-se que haverá um parto, que não é cesariana, ou em Abril ou Junho e por isso contabilizado no segundo trimestre do ano. Para validação desta informação foi verificado que o outro parto contabilizado no segundo trimestre do ano foi realizado em Abril.

Tal como referido, ao seleccionar-se o período de tempo, neste caso, o mês de fevereiro é fornecido mais detalhe sobre esse nível. Esse detalhe pode ser visualizado do lado direito do relatório, onde é indicado o número total de partos, não diferenciando por tipo, e o número de cesarianas. Para este mês ocorreram 164 partos dos quais 40 foram cesarianas, apresentando assim uma taxa de cesarianas de 24,39%.

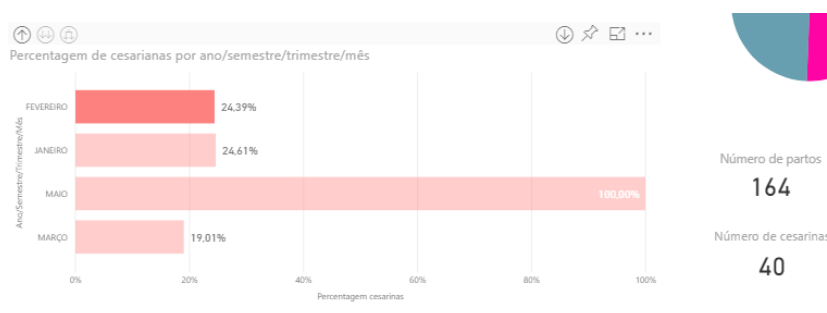


Figura 5.4: Taxa de cesarianas por mês.

Analisando mês a mês, verifica-se que no primeiro trimestre o mês que apresentou uma taxa de cesarianas mais alta foi janeiro, com uma taxa de 24,61%.

Apesar dos valores utilizados serem apenas de teste, servindo apenas para demonstrar a viabilidade da solução proposta é facilmente perceptível os vários benefícios deste relatório para o utilizador final.

Com este relatório o utilizador consegue efetuar o acompanhamento da taxa de cesarianas ao longo do ano, podendo efetuar comparações mês a mês. Num serviço em que o seu financiamento depende desta taxa (Campos et al., 2010) é crucial para um diretor de serviço monitorizar esta taxa, pois assim vai poder tomar medidas ao longo do ano, para que no final o valor desta taxa não comprometa estes financiamentos.

Tal como já referido, na secção 2.5, esta taxa pode ser considerada como um indicador internacional para a monitorização da qualidade dos cuidados prestados no serviço de obstetrícia (WHO, UNFPA e UNICEF, 2009; Long et al., 2015), o fornecimento de informação atualizada e constante auxilia o utilizador na avaliação do seu serviço.

5.1.2 Número de partos realizados

No serviço de obstetrícia são realizados diferentes tipos de parto, nomeadamente, parto eutócico¹ cefálico, parto distócico² parto auxiliado por ventosas, cesarianas, partos vaginais. Para dar resposta aos restantes partos que não sejam cesarianas foi criado um relatório com uma visão geral do número de partos caracterizados pelo seu tipo.

Neste relatório é disponibilizado o filtro do ano, para que o utilizador selecione o período temporal que pretende, sendo posteriormente apresentado para esse ano o número total de partos realizados, e a distribuição desse valor pelos diferentes tipos de partos, quer no gráfico, quer na tabela.

A figura 5.5 apresenta esse relatório, selecionou-se o ano de 2017 sendo apresentado um número total de partos de 479. A tabela apresenta a distribuição desses 479 partos pelos diferentes tipos, é apresentada a contagem para esse tipo de parto, bem como a percentagem que esse parto representa no valor total de partos.

No gráfico circular é apresentada a caracterização dos partos com a sua contabilização para aquele ano. O gráfico apresenta a respetiva legenda no canto inferior esquerdo, sendo que no próprio gráfico são apresentados os diferentes intervalos com a respetiva legenda. Por exemplo, a cor de rosa é representado o número de partos eutócicos cefálicos que são 293.

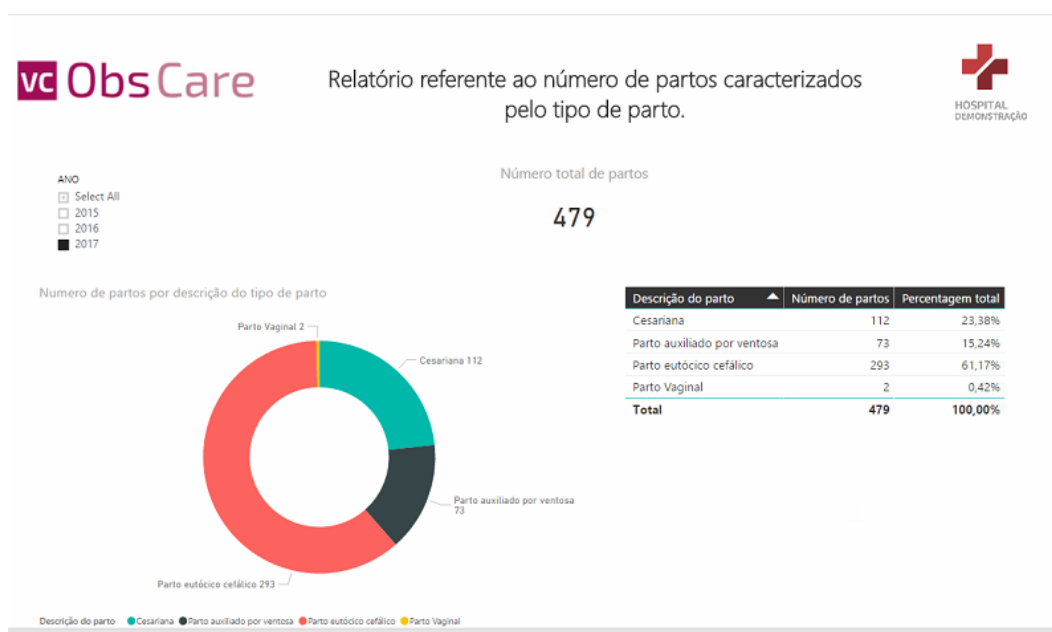


Figura 5.5: Relatório com o número de partos caracterizados pelo tipo de parto.

A disponibilização deste relatório vai permitir ao utilizador acompanhar o número de partos em cada tipo pelos diferentes anos, permitindo uma análise da evolução deste número. O utilizador pode comparar o número de partos eutócicos cefálicos, por exemplo, nos diferentes anos permitindo ao utilizador perceber se esse valor se encontra a subir ou a descer. Assim como, verificar a taxa de cesarianas nos diferentes anos, permitindo verificar essa evolução.

¹Termo equivalente a parto vaginal ou parto normal.

²Parto realizado com recurso a intervenção médica.

Estes relatórios vão permitir ao utilizador obter conhecimento de um determinado conjunto de variáveis, neste caso o número de partos, que vão permitir auxiliar nas decisões tomadas no serviço.

5.1.3 Motivos de uma cesariana

A cesariana é um procedimento cirúrgico, e como tal, tem complicações associadas, e só deve ser realizada quando de facto é necessário. Para monitorização dos motivos que de facto levam a uma cesariana foram definidos motivos principais. Segundo a DGS as cesarianas devem ser classificadas de acordo com esses motivos (George, 2015), que são:

- Patologia materna que contraindica o parto vaginal;
- Anomalia fetal que contraindica o parto vaginal;
- Patologia própria da gravidez;
- Cirurgia uterina prévia;
- Situação ou apresentação fetal anómala;
- Gravidez múltipla;
- Suspeita de incompatibilidade;
- Tentativa frustrada de indução do trabalho de parto;
- Trabalho de parto estacionário; estado fetal não tranquilizador intraparto;
- Outras.

A visualização da contabilização de cesarianas por cada motivo torna-se crucial para o utilizador uma vez que é importante perceber o porquê de se realizar esta cirurgia. Ao entender os motivos e estudar as suas causas é possível tomar medidas que as evitem, o que levará a uma redução da taxa de cesarianas. Desta forma, foi disponibilizado um relatório onde indica o número de cesarianas que ocorrem devido a um determinado motivo.

O relatório apresenta um *dashboard* circular onde apresenta a contabilização das cesarianas por cada motivo, é fornecido também ao utilizador o filtro do ano. O utilizador pode seleccionar o ano que pretende analisar apresentando assim os valores correspondentes a esse ano. A figura 5.6 apresenta o relatório disponibilizado ao utilizador para monitorização destes motivos.

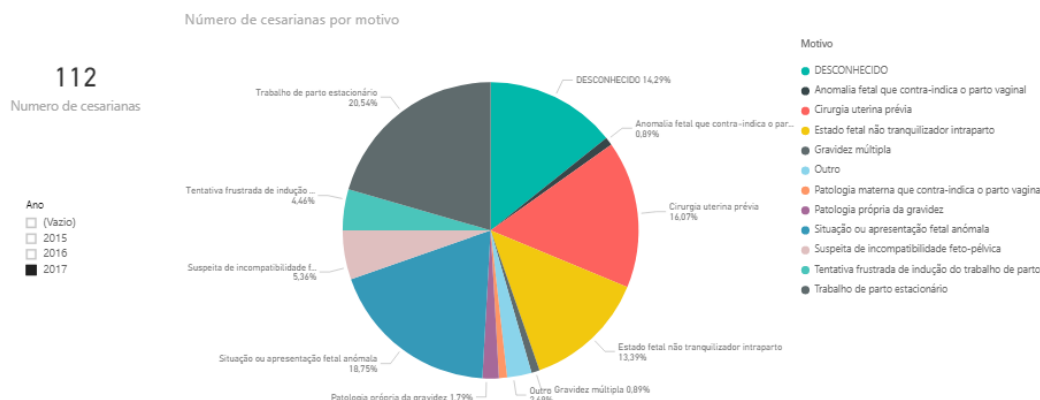


Figura 5.6: Relatório com os motivos das cesarianas.

Tal como nos restantes relatórios o *dashboard* é interativo podendo o utilizador seleccionar um motivo, ou apenas deslizar o rato no *dashboard*, verificando o número de cesarianas para esse motivo e a percentagem que essas cesarianas representam.

Para este caso o ano seleccionado foi 2017, no canto superior esquerdo indica que foram realizadas 112 cesarianas, e através da análise ao *dashboard* verifica-se que o motivo com uma percentagem maior de cesarianas foi o trabalho de parto estacionário (20,54 %), seguindo-se a situação ou apresentação fetal anómala (18,75 %), cirurgia uterina prévia (16,07 %), estado fetal não tranquilizador (13,39 %), suspeita de incompatibilidade pélvica (5,36 %). Para o ano em estudo verificou-se que existiam cesarianas em que não foi atribuído nenhum motivo, de um total de 112, 16 não têm motivo associado. Apesar de os dados utilizados não estarem associados a um hospital, com este relatório o diretor de serviço consegue perceber que existem cesarianas às quais não foi associado um motivo, situação a ocorrer de forma errada.

O utilizador ao verificar quais os principais motivos a ocorrerem no seu hospital vai poder intervir, de forma a tentar minimizar a taxa de partos por cesariana (Campos et al., 2010).

5.1.4 Complicações das cesarianas

São variadas as complicações que podem ocorrer devido a uma cesariana, quer para a mãe, quer para o bebé. Tratando-se de uma cirurgia pode ter complicações anestésicas para a mãe, tendo uma probabilidade de ocorrência duas vezes superior comparativamente com o parto vaginal (Liu et al., 2007).

Outras complicações podem ser lesões nos órgãos; necessidade de transfusões de sangue devido à grande probabilidade de ocorrência de hemorragias; rutura uterina; placenta prévia, accreta e riscos associados podendo levar a uma histerectomia; trombo-embolismo (Mylonas e Friese, 2015a; campos, 2014).

Para além das complicações na mãe, existem as complicações no recém nascido. Uma cesariana efetuada antes das 39 semanas leva a um aumento da morbilidade respiratória no recém nascido, necessitando de cuidados intensivos de saúde, comparativamente com partos vaginais. Outras complicações são a ocorrência de diabetes do tipo 1 na infância, asma, rinite alérgica (Mylonas e Friese, 2015a).

Todas estas complicações levam a que, tanto a mãe como o recém nascido, necessite de mais cuidados médicos, cuidados esses que poderiam ser evitados na ausência de complicações.

Esta monitorização das complicações ocorridas numa cesariana é crucial e para isso foi disponibilizado um relatório que indica o número total de cesarianas, quais delas tiveram complicações e quais são as complicações. Este relatório encontra-se na figura 5.7.



Figura 5.7: Relatório com as complicações das cesarianas.

O ano selecionado para a análise das complicações foi 2017, o número total de cesarianas realizadas nesse ano encontra-se na parte superior, 112. Destas 112, quatro delas tiveram complicações, uma cesariana levou à intubação do recém nascido, em três registou um índice de Apgar ³ <7 ao 5º minuto.

O relatório disponibilizado vai permitir ao utilizador a monitorização destas complicações, podendo efetuar esse acompanhamento por ano.

³Teste que avalia as condições de vitalidade do recém nascido, ou seja, o nível de adaptação à vida extra uterina

5.1.5 Recém nascidos

Depois de se obter um relatório principal focado mais nas taxas de cesariana, foram definidos outros relatórios mais específicos, nomeadamente mais voltados para os recém nascidos.

Foi criado um relatório mais voltado para o número de recém nascidos, onde foram disponibilizados filtros para o utilizador efetuar as suas análises. Os filtros nos relatórios podem ser disponibilizados de duas formas: no *layout* do próprio relatório ou fora.

Assim sendo, foi criado um relatório com o número total de recém nascidos. Para que o utilizador possa efetuar a manipulação dos dados foram fornecidos dois filtros no *layout* do relatório que são: o ano de nascimento do bebé e o tipo de parto. O utilizador ao filtrar os dados por essas variáveis, visualiza apenas o número de recém nascidos pelo tipo de parto escolhido e para o ano também selecionado. Em paralelo é fornecido outro filtro ao utilizador, com o sexo do recém nascido (F - feminino, M - masculino). Esta variável é fornecida fora do *layout* do relatório, no entanto, o filtro afeta todos os dados daquela página do relatório.

Neste relatório, figura 5.8, são apresentados também o número de recém nascidos com mal formações e sem mal formações, *dashboard* à direita na parte inferior do relatório, e o número de recém nascidos vivos e mortos, *dashboard* à esquerda.

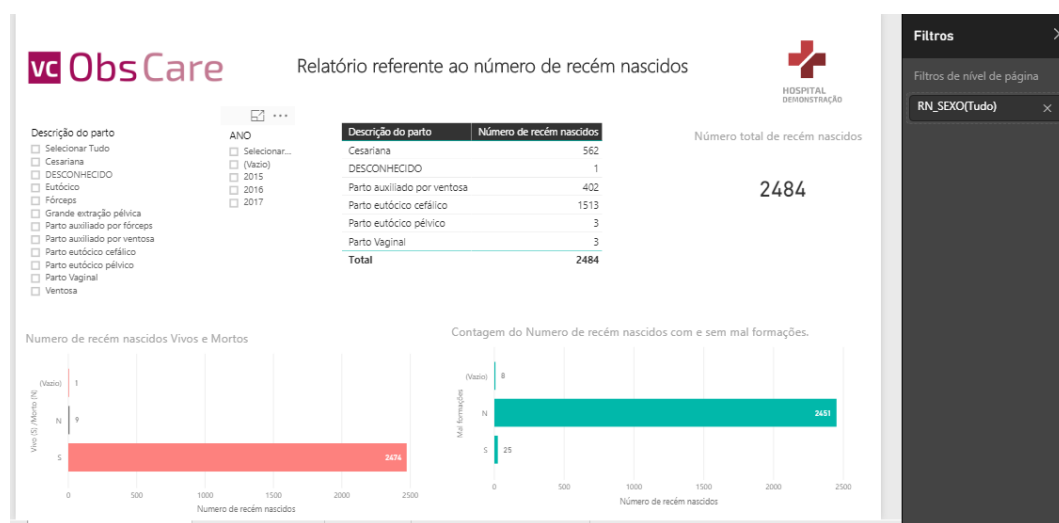


Figura 5.8: Relatório com o número de recém nascidos.

Com a disponibilização dos filtros o utilizador pode filtrar pelos campos que pretende podendo efetuar uma análise mais personalizada e detalhada. Por exemplo, se o utilizador pretender visualizar o número de recém nascidos referentes a 2017 do sexo masculino, pode aplicar o filtro do ano no *layout* do relatório e aplicar o filtro do sexo que se encontra fora do relatório, figura 5.9. Sendo que os dois filtros afetam todos os dados presentes naquela página. Os dados apresentados na tabela e nos *dashboards* são apenas referentes aos filtros aplicados, assim como o valor apresentado para o número total de recém nascidos.

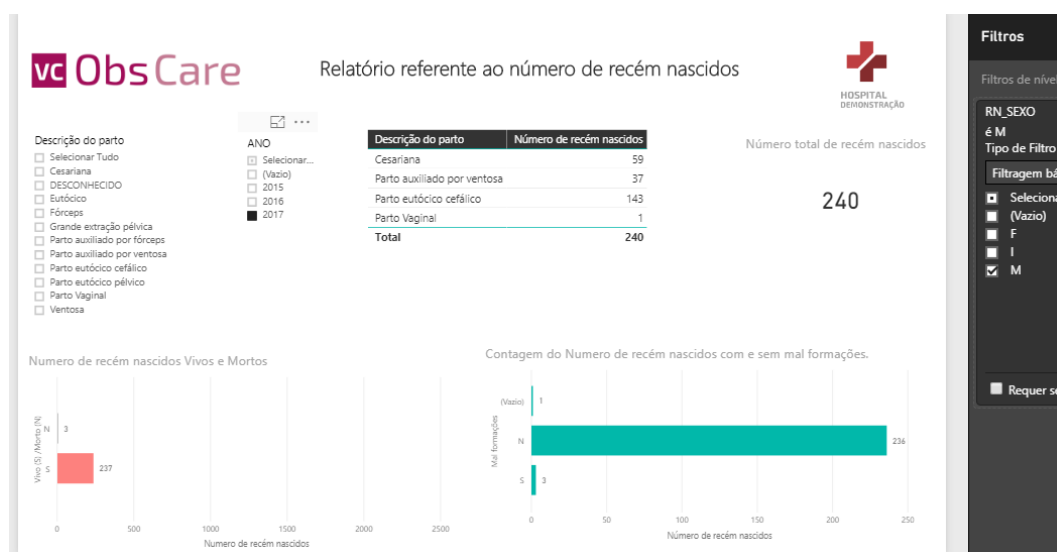


Figura 5.9: Relatório com o número de recém nascidos do ano 2017, do sexo masculino.

Neste relatório o número de recém nascidos pode ser filtrado com recurso a três filtros: tipo de parto; ano; sexo do recém nascido. Para além destes filtros o número de recém nascido pode ainda ser dividido por dados do bebé: apresentação de mal formações; se nasceu vivo ou morto.

Esta filtragem e subdivisão dos dados vai permitir ao utilizador uma interação com os dados permitindo-lhe efetuar análises mais detalhadas.

Uma das variáveis que o utilizador pode avaliar é a natalidade (número de recém nascidos que nasceram vivos) naquele serviço para aquele ano. O mesmo *dashboard*, no canto inferior esquerdo, apresenta também o número de recém nascidos que nasceram mortos, o que pode permitir uma análise da mortalidade para aquele período de tempo dos recém nascidos.

Se a mortalidade dos recém nascidos apresentar um valor em crescimento, por exemplo, o utilizador pode tentar perceber o porque desse valor, se de facto esses valores se devem a causas naturais ou se por algum motivo não estão a ser prestados os cuidados corretos e necessários durante o nascimento do recém nascido.

Em conjunto com o cliente percebeu-se também que uma das suas necessidades é a análise da distribuição do número de nascimentos durante o ano, tendo-se a necessidade de chegar a cada mês desse ano. Desta forma o número de recém nascidos distribuído ao longo do tempo também se torna numa variável importante para análise, assim sendo, foi disponibilizado um relatório com dois *dashboards*. Um deles apresenta o número de recém nascidos por ano, no outro é disponibilizado o número de recém nascidos por mês, podendo-se perceber em que mês ocorreram mais nascimentos. A figura 5.10 apresenta o relatório disponibilizado com estas métricas.

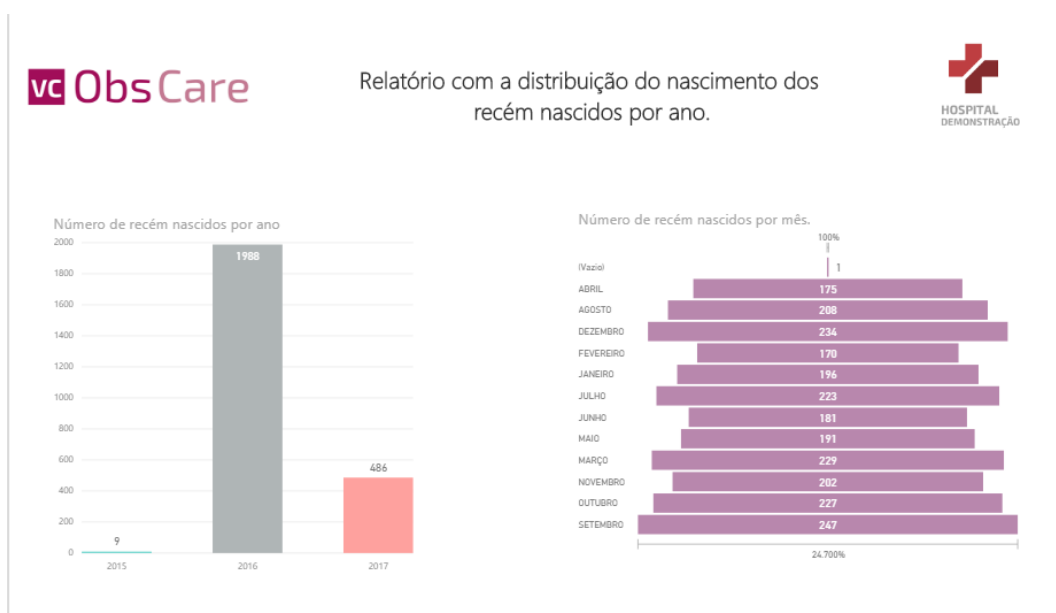


Figura 5.10: Relatório com o número de recém nascidos distribuídos por ano e por mês.

Neste relatório não é disponibilizado nenhum filtro, no entanto, como os *dashboards* são interativos consegue-se manipular os dados selecionando o ano ou mês que se pretende.

O utilizador pode selecionar o ano correspondente ao período que pretende analisar e o *dashboard* com os meses vai apresentar a distribuição dos recém nascidos daquele ano pelos meses correspondentes, figura 5.11.

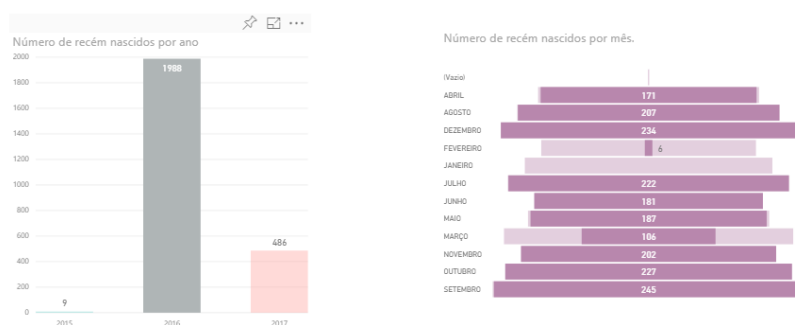


Figura 5.11: Filtragem da distribuição de recém nascidos por mês para o ano de 2016.

O contrário também se verifica, isto é, se o utilizador selecionar um mês, será indicado no gráfico dos valores anuais, a distribuição do número de recém nascidos que nasceram naquele mês nos respetivos anos. Na figura 5.12 foi selecionado o mês de janeiro, o que indica que nasceram 196 recém nascidos apenas no ano de 2017, nos anos anteriores, 2015 e 2016 não se registaram nascimentos para esse mês.

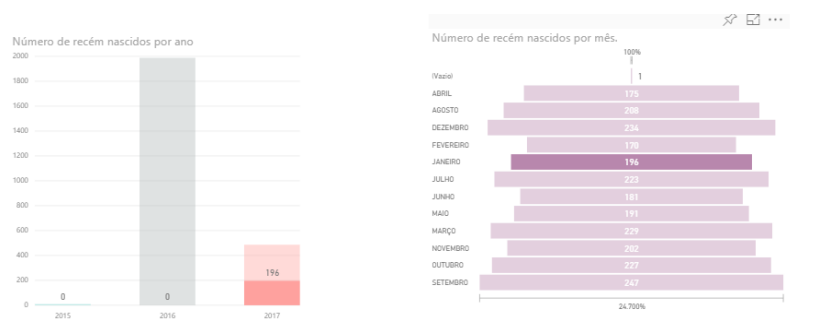


Figura 5.12: Filtragem da distribuição de recém nascidos para o mês de janeiro, por ano.

Com a disponibilização deste relatório o utilizador vai poder analisar os dados até ao mês, podendo mesmo verificar para um determinado ano em que mês ocorreram mais nascimentos.

Outro fator importante, que o cliente referiu, é o relacionamento da posição do bebé nas cesarianas. Variável também já encontrada pela literatura em que esta apresentação do bebé pode ser considerada, por algumas evidências, como um dos motivos para efetuar uma cesariana (Campos et al., 2010).

Aliando a apresentação do bebé com o seu peso foi disponibilizado um relatório com estas variáveis. Neste relatório consta uma tabela que indica a posição do bebé e o número de cesarianas realizadas com o bebé naquela apresentação, é também fornecido um *dashboard* com o número de cesarianas caracterizadas pelo peso do recém nascido.

A variável do peso foi dividida em intervalos, por exemplo para o intervalo [2500;3000] são contabilizadas as cesarianas com bebés com um peso superior a 2500 gramas e inferior ou igual a 3000 gramas.

A figura 5.13 apresenta este relatório, tal como os restantes o *dashboard* que apresenta os níveis do peso é interativo e se o utilizador selecionar um intervalo de peso é indicado na tabela em que posição se encontravam os recém nascidos desse intervalo.

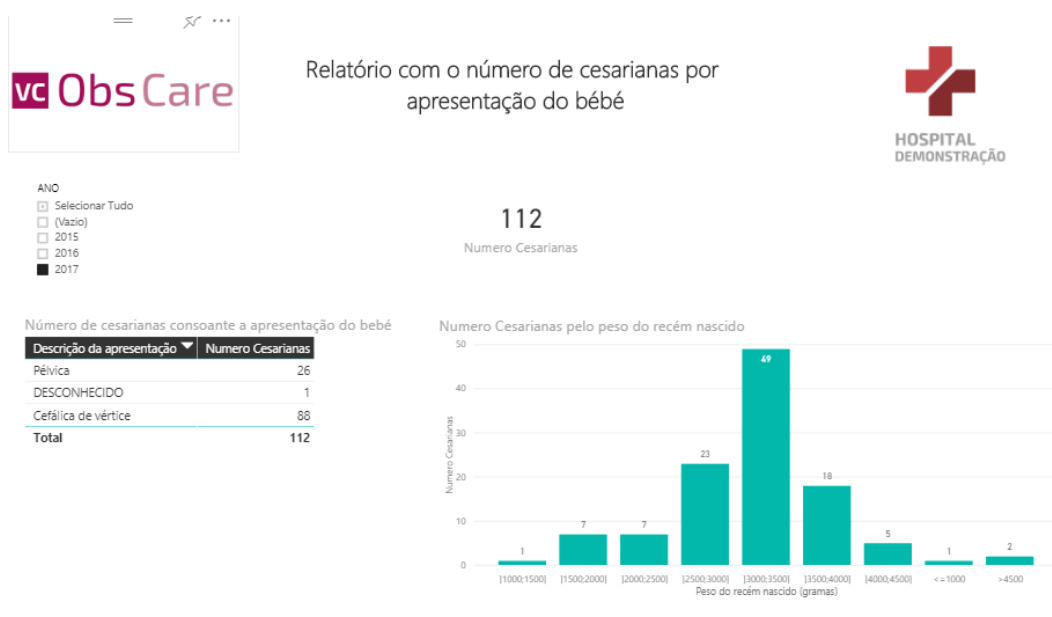


Figura 5.13: Relatório com o número de cesarianas por apresentação do bebé.

Um dos motivos para a ocorrência de uma cesariana, nos hospitais públicos da região norte, é a incompatibilidade feto-pélvica (Campos et al., 2010), isto é, o bebé é demasiado grande para a estrutura pélvica da mulher. Com este relatório o utilizador vai poder incidir algumas análises relativamente à apresentação do bebé e ao peso, sendo que, é possível selecionar um dos níveis de peso do recém nascido e verificar de que forma esses bebés se apresentavam.

Outro motivo para efetuar uma cesariana é a apresentação pélvica do bebé, sendo que quando o bebé não se encontra numa posição favorável ao parto vaginal pode ser tentada uma manobra para manipulação dessa apresentação. Esta técnica incide em manipular o bebé exteriormente de forma a virá-lo, sendo efetuada sob vigilância ecográfica e requerendo alguma experiência por parte do médico (Campos et al., 2010). Se não existir nenhuma contra-indicação é uma técnica recomendável, sendo que os médicos têm de ter conhecimentos específicos para implementação desta manobra.

A informação fornecida neste relatório vai permitir ao diretor de serviço verificar se o número de cesarianas devido à apresentação do bebé é elevado, e caso isso se verifique, pode tomar medidas para que seja tentada a versão cefálica externa com mais frequência e para isso o hospital deve fornecer constantemente cursos regulares de treino para os médicos na execução da versão cefálica externa (Campos et al., 2010).

5.1.6 Caracterização da mãe em cesarianas

Outros fatores importantes são as características da mãe (idade e peso). Como forma de relacionar estas características com o número de cesarianas, foi elaborado um relatório. O relatório visualizado na figura 5.14 disponibiliza dois *dashboards*, à esquerda é indicado o número de cesarianas caracterizadas pelo peso da mãe, e à direita é apresentado o número de cesarianas caracterizado pela idade da mãe.

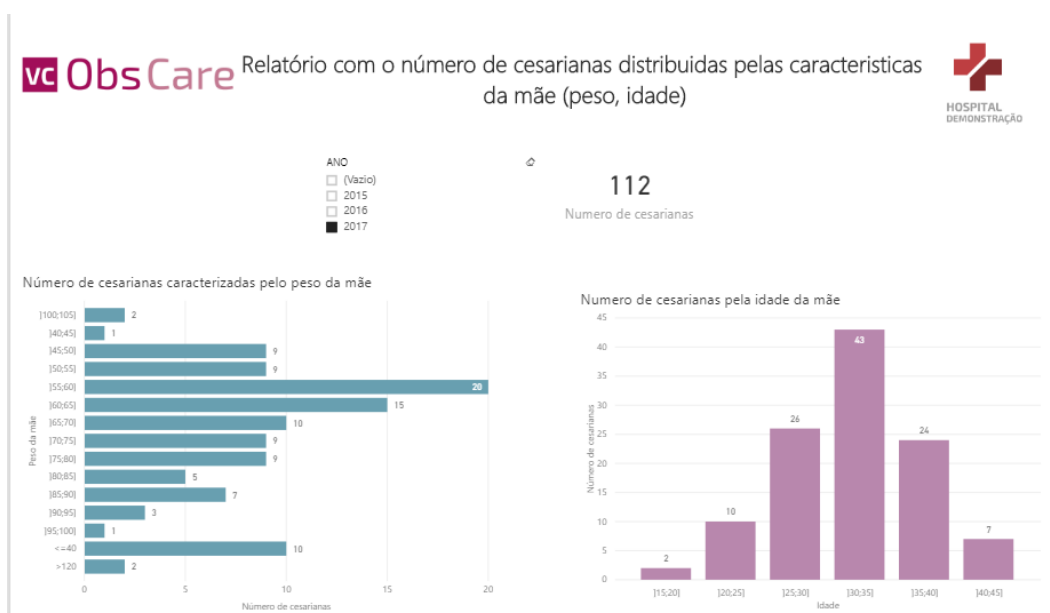


Figura 5.14: Relatório do número de cesarianas caracterizadas pelo peso e idade da mãe.

A idade materna parece ter um papel importante na decisão por uma cesariana, uma vez que cada vez mais a maternidade é adiada levando a que a média da idade materna aumente. Quanto mais alta é esta idade mais complicações podem ocorrer, sendo que, gestações ocorridas em mulheres com uma idade superior a 35 anos podem ser consideradas como alto risco, como há uma maior probabilidade de ocorrerem riscos no parto na maior parte das vezes opta-se pela cesariana para minimizar esses riscos (Mylonas e Friese, 2015b).

Outras características da mãe, como algumas doenças, aumentam a probabilidade de ocorrerem riscos, por exemplo o excesso de peso na mãe pode levar à hipertensão (Mylonas e Friese, 2015b; Khalil et al., 2013). O conjunto de riscos que podem surgir no parto leva a que a decisão pela cesariana seja considerada preferencialmente. Com o relatório fornecido ao utilizador vai ser possível avaliar estas características bem como esta predisposição para a escolha pela cesariana.

5.1.7 Tipo de episódios

Outro indicador referenciado pelo cliente e também obtido atualmente é referente à contabilização dos episódios desse serviço caracterizados pelo seu tipo, isto é, quantos episódios ocorreram do tipo consulta, por exemplo.

Para dar resposta a este indicador foi criado um relatório, figura 5.15 onde é apresentado esse indicador, numa forma mais gráfica, lado direito do relatório e também numa tabela, lado esquerdo.

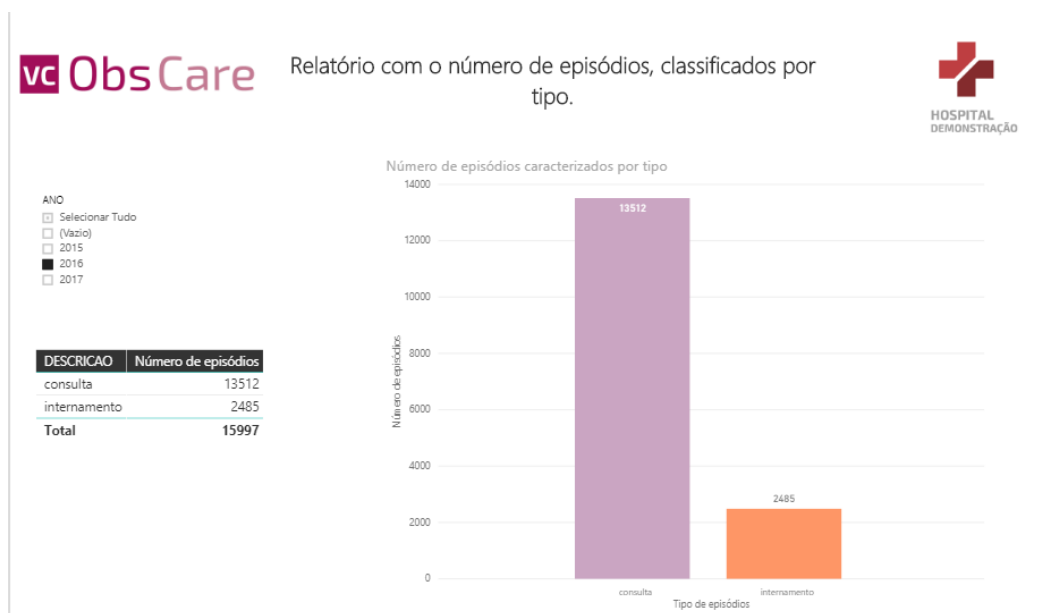


Figura 5.15: Relatório com o número de episódios diferenciados pelo seu tipo.

Para que o utilizador possa interagir com os resultados e analisar apenas o período de tempo que pretende, o ano, é fornecido um filtro. No relatório apresentado foi selecionado o ano de 2016 apresentando 13512 episódios de consulta e 2485 episódios de internamento naquele hospital.

Este indicador, dá informação da utilização do serviço, contabilizando o número de internamentos, consultas, e urgências efetuadas naquele serviço.

5.1.8 Comparação entre instituições

A taxa de cesarianas é considerada um indicador internacional para a caracterização dos cuidados obstétricos (WHO, UNFPA e UNICEF, 2009; Long et al., 2015). Em Portugal o valor dessa taxa ainda não é o desejável, sendo o último valor conhecido referente a 2015, com um valor provisório de 32,5 % (INE, DGS MS e PORDATA, 2016).

Para que se possa analisar hospital a hospital é importante ter dados referentes a todos ele de forma a que possam ser comparados. Para a ARS do norte é importante efetuar esta comparação entre instituições, uma vez que nelas têm instalado o mesmo sistema de recolha de dados, ObsCare, sendo comparável a taxa de cesarianas de uma instituição com a outra.

Para dar resposta a esta necessidade foi disponibilizado um relatório com a taxa de cesarianas de cada instituição. No relatório é possível comparar essa taxa por ano, sendo também disponibilizado a lista de hospitais disponíveis para análise.

O relatório disponibilizado para esta comparação, encontra-se na figura 5.16, onde foram utilizados dados referentes a duas instituições de teste.

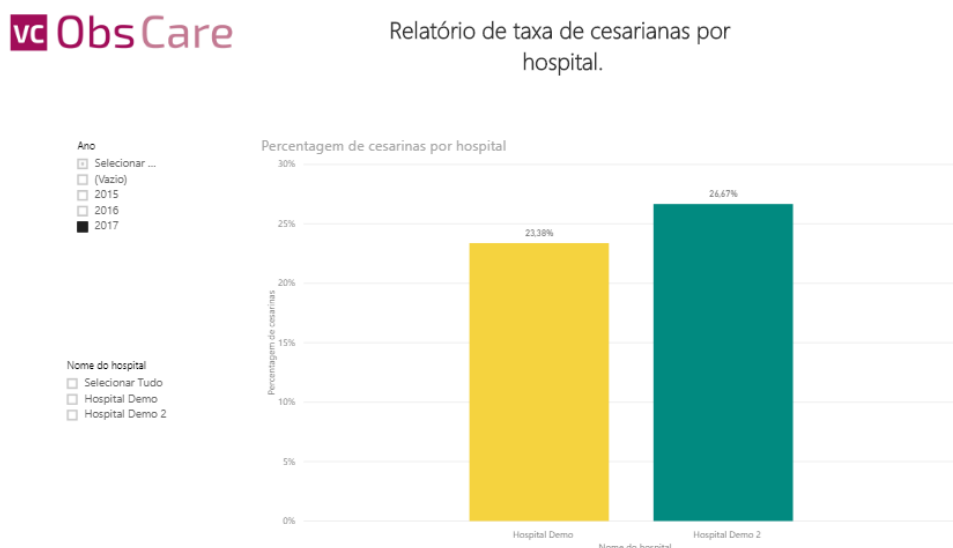


Figura 5.16: Relatório com a taxa de cesarianas dos diferentes hospitais.

Para o ano selecionado, 2017, um hospital apresentaria uma taxa de 23,38 % (instituição sob a qual incidiram todas as outras análises) e outro hospital apresentaria uma taxa de 26,67 %. As taxas de cada instituição são facilmente perceptíveis pelo *dashboard* à direita, percebendo-se visualmente qual o hospital que apresenta uma taxa superior, no caso em estudo, o hospital demo 2 apresenta uma taxa de cesarianas superior.

Neste caso, o utilizador seria a ARS responsável tendo acesso às taxas de todos os seus hospitais. A ARS poderia definir medidas para serem implementadas em todos os hospitais e com este relatório verificar o impacto dessas medidas em cada hospital.

5.2 Conclusão

A ferramenta desenvolvida centraliza a informação de diferentes hospitais, o processamento de todos esses dados permitiu dar resposta aos requisitos definidos.

Os requisitos definidos para a solução foram:

- Filtrar;
- Executar relatórios;
- Visualizar *dashboards*;
- Comparação interinstitucional.

A filtragem da informação é garantida pelo fornecimento de filtros no *layout* do relatório, na maioria dos relatórios o filtro fornecido corresponde ao ano para que o utilizador possa analisar um ano específico.

Em todos os relatórios são fornecidos *dashboards* em que é possível ao utilizador interagir com a informação, podendo mesmo efetuar filtragens ao selecionar determinados dados no *dashboard*.

Os relatórios podem ser guardados em diferentes formatos, podendo também imprimir-se diretamente da aplicação, figura 5.17.

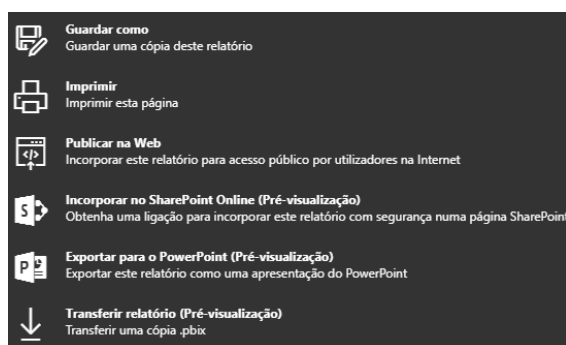


Figura 5.17: Apresentação das opções de execução do relatório fornecidas pelo Power BI.

Outro requisito definido para a ferramenta desenvolvida foi a disponibilização de comparações entre hospitais. Esta funcionalidade é garantida pelo relatório que apresenta a taxa de cesarianas dos diferentes hospitais. De salientar que esta funcionalidade é apenas para os profissionais da ARS o único utilizador que visualiza a informação de todos os hospitais. O diretor de serviço de um hospital apenas tem acesso aos dados do seu hospital.

Capítulo 6

Avaliação da Solução

Neste capítulo é apresentada a avaliação da solução desenvolvida para a análise dos dados do setor de obstetrícia. A solução foi avaliada através de inquéritos de satisfação e usabilidade aos utilizadores da ferramenta e aos colaboradores da VirtualCare.

6.1 Metodologia de avaliação

A avaliação da solução apresentada foi efetuada com recurso a inquéritos de satisfação e de usabilidade. Para a elaboração destes inquéritos foram definidos dois grupos distintos de profissionais: os médicos e os colaboradores da VirtualCare.

Os médicos escolhidos pertencem a hospitais que já possuem o ObsCare nos seus serviços. Estes profissionais fazem parte do serviço de obstetrícia dos vários hospitais onde atualmente a disponibilização dos indicadores é efetuada de forma estática. Assim sendo foi utilizada uma amostra por conveniência uma vez que são profissionais ligados diretamente ao serviço sabendo das necessidades deste, tendo uma maior capacidade de inferir a utilidade desta solução e do que pode ser melhorado. Estes profissionais podem também avaliar a solução com base na solução que utilizam diariamente, podendo assim efetuar comparações entre as ferramentas.

Foram elaborados dois inquéritos distintos, um para os médicos outro para os colaboradores da VirtualCare, uma vez que cada grupo vai avaliar características diferentes da solução desenvolvida.

Cada inquérito possuía uma escala de respostas representadas por números para facilitar a resposta e a análise destas respostas. A escala disponibilizada nos dois inquéritos encontra-se na tabela 6.1.

Tabela 6.1: Escala utilizada nos inquéritos.

	Escala
Discordo completamente	1
Discordo	2
Não tenho opinião	3
Concordo	4
Concordo completamente	5

O inquérito disponibilizado aos médicos, anexo B, apresenta comentários relativamente à forma com esta nova solução se adequa ou não ao serviço, bem como, se ocorreram melhorias com esta nova abordagem. Para que os médicos pudessem avaliar a solução foi disponibilizado um vídeo onde eram apresentadas de uma forma genérica as funcionalidades da nova solução de BI para o serviço de obstetria.

O inquérito para os colaboradores da VirtualCare, anexo C, apresenta comentários mais técnicos, mais direcionados para a arquitetura desenvolvida e como esta pode ou não ser integrada com outros sistemas.

6.2 Resultados do inquérito disponibilizado aos médicos

O inquérito disponibilizado aos médicos apresenta oito afirmações com resposta para a escala apresentada anteriormente e uma pergunta com resposta aberta. O inquérito foi disponibilizado a 50 médicos de diferentes instituições de saúde, sendo que 25 % destes responderam ao inquérito.

A tabela 6.2 apresenta a frequência relativa para cada valor da escala de respostas das várias afirmações do inquérito. A identificação destas afirmações encontram-se assinaladas com a letra A, seguindo-se da numeração de cada uma. Na oitava pergunta um dos utilizadores não selecionou nenhuma das opções fornecidas, daí não se obter uma percentagem de 100% na resposta desta questão.

Tabela 6.2: Frequência relativa de cada resposta do inquérito disponibilizado aos médicos.

	1	2	3	4	5
A1	0%	0%	11,1%	44,4%	44,4%
A2	0%	0%	11,1%	77,8%	11,1%
A3	0%	0%	11,1%	33,3%	55,6%
A4	0%	0%	11,1%	55,6%	33,3%
A5	0%	0%	33,3%	44,4%	22,2%
A6	0%	0%	0%	44,4%	55,6%
A7	0%	0%	0%	22,2%	77,8%
A8	0%	0%	0%	0%	88,9%

Após a análise da tabela verifica-se que para as quatro primeiras afirmações, aproximadamente 90% das respostas incidiram sobre o 4 (concordo) e o 5 (concordo completamente),

os restantes utilizadores não tiveram opinião sobre estas afirmações. No que diz respeito à utilização da aplicação, 44,4% dos utilizadores concordava e 44,4% concordava completamente que a solução era fácil de utilizar sendo intuitiva e simples. Relativamente à pertinência dos filtros disponibilizados, 77,8% dos utilizadores concordou que os filtros apresentados eram pertinentes para a realidade do hospital, 11,1% dos utilizadores concordou completamente com esta afirmação. Quanto ao tempo de resposta da solução, 33,3% dos utilizadores concordou que a aplicação respondia rapidamente quando se aplicavam filtros, 55,6% concordou completamente. Relativamente à disponibilização da informação segundo esta abordagem, 55,6% dos utilizadores concordaram que a disponibilização segundo esta abordagem vai permitir analisar de forma mais eficiente a taxa de cesarianas, 33,3% concordaram completamente com esta afirmação.

Quando o utilizador é abordado sobre a adequação dos indicadores fornecidos ao hospital (afirmação cinco), 33,3% indicaram que não tinham opinião, sendo que 44,4% concordaram com os indicadores disponibilizados e 22,2% concordaram completamente.

Nas restantes afirmações, as respostas incidiram apenas sobre o 4 (concordo) e o 5 (concordo completamente). Os utilizadores preferiram a apresentação dos indicadores sobre esta forma, 44,4% dos utilizadores concordou que prefere esta forma de apresentação de indicadores ao invés de relatórios estáticos e 55,6% dos utilizadores concorda completamente com esta disponibilização dos indicadores. Relativamente às melhorias encontradas nesta abordagem, 22,2% concordou que comparativamente com a solução atual dos serviços, existem melhorias com a utilização desta solução, 77,8% concordou completamente. Quando os utilizadores foram abordados sobre a possibilidade de ter esta solução no seu serviço, todos eles demonstraram interesse em adquirir esta nova solução para o seu hospital.

Na pergunta de resposta aberta, onde são questionados se acrescentariam mais algum indicador à ferramenta, e qual, não foram fornecidas nenhuma sugestões de novos indicadores, o que indica que, de uma forma geral os utilizadores ficaram satisfeitos com os indicadores disponibilizados.

6.3 Resultados do inquérito disponibilizado aos colaboradores da VirtualCare

O inquérito elaborado para os colaboradores da VirtualCare apresenta quatro afirmações técnicas, mais voltadas para a arquitetura da solução. Para este inquérito foram obtidas duas respostas.

A tabela 6.3 apresenta a frequência relativa das respostas para cada pergunta apresentada.

Tabela 6.3: Frequência relativa de cada resposta do inquérito aos colaboradores da VirtualCare.

	1	2	3	4	5
A1	0%	0%	0%	100%	0%
A2	0%	0%	0%	0%	100%
A3	0%	0%	50%	50%	0%
A4	0%	0%	0%	0%	100%

As respostas às duas primeiras questões incidiram sobre a resposta 4 (concordo) e 5 (concordo completamente). Todos os inquiridos, 100%, concordaram que a arquitetura da solução seria adequada às necessidades de negócio. Quando abordados sobre se a arquitetura iria permitir crescer para novos indicadores, todos os inquiridos concordaram completamente que a solução permite crescer para novos indicadores.

Na terceira afirmação, a resposta dos inquiridos foi dividida entre o 3 (não tem opinião) e 4 (concordo). Metade dos utilizadores não tinham opinião e a outra metade concordou que a aplicação seria facilmente integrada com outros sistemas.

A última resposta foi unânime, todos os inquiridos responderam 5 (concordo completamente), o que indica que todos eles concordaram que esta nova solução trouxe melhorias comparativamente com a solução atual.

6.4 Conclusão

Os inquéritos disponibilizados permitiram avaliar se a solução correspondia às necessidades do serviço de obstetrícia, bem como, se seria importante no quotidiano dos médicos podendo-os ajudar, e dar a informação necessária, para analisarem os indicadores de uma forma mais eficiente.

Com base em todas as respostas obtidas, apesar de estas respostas terem sido efetuadas por uma amostra relativamente pequena, pode-se concluir que os resultados foram positivos apresentando um grau de satisfação elevado por parte dos inquiridos.

A solução apresentou também uma elevada percentagem de aceitação, uma vez que, das pessoas inquiridas todas elas mostraram interesse em adquirir a nova solução.

Como a amostra de utilizadores que conseguiu responder é pequena, numa fase posterior estes inquéritos deveriam ser efetuados novamente para uma validação mais consistente da aplicação. Nesta nova disponibilização dos inquéritos deverá alargar-se o tempo de resposta, porque este pode ter sido a causa do número reduzido de respostas.

Capítulo 7

Conclusão

A elaboração desta dissertação teve como principal objetivo o desenvolvimento de um sistema de informação de BI capaz de fornecer informação necessária aos profissionais de saúde. Este sistema é direcionado para os profissionais dos serviços de obstetrícia sendo lhes fornecido um conjunto de informação que os vai apoiar no processo de redução da taxa de cesarianas melhorando assim a qualidade deste serviço. De forma a atingir este objetivo foi utilizado como caso de estudo o serviço de obstetrícia do Centro Hospitalar de São João.

Numa primeira fase foi elaborada uma análise que permitisse contextualizar o problema abordado na dissertação originando uma proposta de valor e uma análise do contributo das aplicações de BI no âmbito da saúde. Foram ainda abordados os principais conceitos e a aplicabilidade de uma aplicação de BI num serviço de obstetrícia. Após esta fase foi possível concluir que este tipo de aplicação apresenta um enorme conjunto de vantagens quando aplicada na saúde, de entre as quais se destaca a rapidez com que é possível apresentar a informação ao profissional de saúde permitindo que este disponha de mais tempo para desempenhar as restantes funções.

Com uma análise detalhada encontraram-se algumas aplicações que respondiam ao problema abordado nesta dissertação. A avaliação destas aplicações foi efetuada com a utilização de um método de apoio à decisão AHP e foi possível verificar que o Power BI é a ferramenta de BI que mais se adequa às necessidades. Esta aplicação apresenta um método de configuração mais intuitivo, uma melhor performance e maior facilidade de integração quando comparada com as restantes.

No processo de decisão de qual a melhor arquitetura a implementar foi escolhida a arquitetura que centraliza todos os dados num modelo dimensional. Esta arquitetura permite a centralização dos dados referentes a diferentes hospitais e armazenar histórico, não comprometendo a rapidez da apresentação de dados ao utilizador. A arquitetura definida permite ainda assegurar os requisitos funcionais de filtragem de informação, execução de relatórios, apresentação dos mesmos em *dashboards* e comparação da taxa de cesarianas entre hospitais.

De forma a fornecer informação que permita melhorar a qualidade de um serviço de obstetrícia foram definidos e implementados os seguintes indicadores:

- Número de cesarianas anual;
- Número de partos realizados;
- Motivos de uma cesariana;
- Número de recém nascidos;

- Relação da apresentação do bebê com as cesarianas;
- Características da mãe nas cesarianas;
- Tipo de episódios;
- Taxa de cesarianas por hospital.

A implementação destes indicadores permite para além de calcular a taxa, compreender qual a sua causa. A disponibilização dos motivos de uma cesariana permite apresentar os principais motivos/causas que originaram uma cesariana. As características da mãe, bem como a apresentação do bebê, procuram tirar algumas conclusões de como estas características podem levar à ocorrência de cesarianas.

O indicador referente às complicações de cesariana permite compreender que uma cesariana para além de ser uma cirurgia e apresentar complicações provenientes da cirurgia, possui também outros riscos para a mãe e para o bebê.

Enquanto os outros indicadores permitem ter uma visão do panorama geral do número de partos e da incidência das cesarianas quando comparadas com os partos vaginais, o indicador com a comparação da taxa de cesarianas por hospital é específico a cada local e estabelece uma relação entre a taxa de cesarianas dos vários hospitais.

Todos os indicadores disponibilizados não utilizam dados nominais dos utentes, desta forma, o carregamento dos dados pode ser efetuado de forma anonimizada, carregando-se apenas atributos como data de nascimento, sexo entre outros. Esta anonimização de dados vai de encontro à nova lei de proteção de dados (Europeia, 2016), em que o utilizador tem o poder sobre os mesmos. Carregando-se os dados desta forma não vai ocorrer nenhuma implicação na solução desenvolvida.

Com a implementação destes indicadores, cumprindo os requisitos funcionais anteriormente referidos, é possível afirmar que o objetivo desta dissertação foi cumprido. De forma a comprovar o cumprimento destes objetivos foram realizados inquéritos de satisfação e usabilidade a uma amostra por conveniência (profissionais de saúde do serviço de obstetria e técnicos de informática).

Os inquéritos revelam um elevado grau de satisfação no tempo de resposta da aplicação, no cumprimento dos requisitos funcionais e nos indicadores implementados.

Conclui-se assim que o objetivo da dissertação foi cumprido com sucesso e é possível com a implementação de um sistema de BI apoiar na compreensão da taxa de cesarianas e por consequência a sua redução.

7.1 Desenvolvimentos Futuros

Atualmente os sistemas de informação desempenham um papel muito importante em qualquer área, destacando-se a área da saúde pelo seu impacto no quotidiano da população. Como tal, existe uma crescente necessidade de desenvolver sistemas de informação de forma a solucionar os problemas apresentados nesta área.

No caso desta dissertação é ainda necessário desenvolverem-se mais indicadores que procurem justificar os elevados valores das taxas de cesariana, relacionando-os com as características da mãe e do bebê de forma a que situações destas possam ser evitadas ou previstas.

Indicadores como percentagem de 1º parto por cesariana ou taxa de partos vaginais após a primeira cesariana, não foram obtidos nesta fase do protótipo, no entanto são importantes para compreender o panorama de cesarianas e devem ser obtidos numa fase posterior.

Numa fase futura pode também efetuar-se a comparação de todos os indicadores entre as instituições de saúde e não apenas a taxa de cesarianas. Esta comparação é extremamente importante para que se possa compreender as principais diferenças entre estas instituições e de que forma essas diferenças podem afetar os diferentes indicadores.

A partilha do relatório gerado pelo Power BI é ainda um assunto que carece de investigação mais cuidada de forma a averiguar se existem mais opções e potencialmente mais adequadas ao contexto apresentado.

É fundamental o teste dos indicadores em ambiente real de forma a que seja possível tirar conclusões definitivas das principais causas da taxa de cesariana e assim desenvolver um plano que permita reduzi-la.

Outro aspeto a ser analisado seria a possibilidade da ferramenta prever a evolução da taxa de cesarianas de acordo com os indicadores que a influenciam.

Bibliografia

- Dale, Sanders (2014). «It All Starts with a Data Warehouse». Em: p. 14.
- Peng, Thomas C C et al. (1992). «An Integrated, Hospital Information System Based Obstetrical Medical Record and Database». Em: url: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2247601/pdf/procascamc00004-0613.pdf>.
- Mylonas, Ioannis e Klaus Friese (2015a). «Indications for and Risks of Elective Cesarean Section.» Em: *Deutsches Arzteblatt international* 112.29-30, pp. 489–95. issn: 1866-0452. doi: 10.3238/arztebl.2015.0489. url: <http://www.ncbi.nlm.nih.gov/pubmed/26249251>.
- VirtualCare, Lda. *VC ObsCare – VirtualCare – Systems for Life*. url: <http://virtualcare.pt/index.php/portfolio/vc-obscare/> (acedido em 21/09/2017).
- George, Francisco Henrique Moura (2015). *Norma nº 001/2015 de 19/01/2015 - Registo de Indicações de Cesariana*. Lisboa.
- Hannah Da Silva Copelli, Fernanda et al. (2015). «Fatores determinantes para a preferência da mulher pela cesariana.» Em: *Abr-Jun* 24.2, pp. 336–43. doi: 10.1590/0104-07072015000430014.
- Faisal-Cury, Alexandre e Paulo Rossi Menezes (2006). «Fatores associados à preferência por cesareana». Em: *Revista de Saúde Pública* 40.2, pp. 226–232. issn: 0034-8910. doi: 10.1590/S0034-89102006000200007. url: http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102006000200007&lng=pt&nrm=iso&tlng=pt.
- Mettler, Tobias e Vivian Vimarlund (2007). «Understanding business intelligence in the context of healthcare». Em: 15153.3. doi: 10.1177/1460458209337446. url: <http://www.sagepub.co.uk/journalsPermissions.navwww.sagepublications.com>.
- Ashrafi, Noushin, Lori Kelleher e Jean-Pierre Kuilboer (2014). «The Impact of Business Intelligence on Healthcare Delivery in the USA». Em: *Interdisciplinary Journal of Information* 9.9, pp. 117–130. issn: 15551237. doi: 10.1016/S2212-5671(14)00859-4. url: <http://www.ijikm.org/Volume9/IJIKMv9p117-130Ashrafi0761.pdf>.
- Raghupathi, W e V Raghupathi (2013). «An Overview of Health Analytics». Em: *J Health Med Informat* 4.3. issn: 21577420. doi: 10.4172/2157-7420.1000132.
- Direção-Geral da Saúde (2016). *Indicadores de Qualidade*. url: <http://www.dgs.pt/qualidade-e-seguranca/monitorizacao/indicadores-de-qualidade1.aspx> (acedido em 27/02/2017).
- INE, DGS MS e PORDATA (2016). *PORDATA - Cesarianas nos hospitais (%) em Portugal*. url: [http://www.pordata.pt/Portugal/Cesarianas+nos+hospitais+\(percentagem\)-1985](http://www.pordata.pt/Portugal/Cesarianas+nos+hospitais+(percentagem)-1985) (acedido em 31/01/2017).
- INE, DGS/MS e PORDATA (2017). *PORDATA - Partos: total e em estabelecimentos de saúde - Portugal*. url: <http://www.pordata.pt/Portugal/Partos+total+e+em+estabelecimentos+de+saude-152> (acedido em 07/02/2017).
- Campos, Diogo Ayres et al. (2010). «Medidas para reduzir a taxa de cesarianas na região norte de Portugal». Em: *Comissão para a redução da taxa de cesarianas da ars norte*, pp. 1–18.

- Methods, Effective et al. (1996). «Fuzzy Front End : and Techniques». Em: *Industrial Research*.
- Administração Regional de Saúde do Norte, I.P. (2017). *Hospitais*. url: <http://portal.arsnorte.min-saude.pt/portal/page/portal/ARSNorte/Institui{\c{c}}{\~{o}}\es/Hospitais> (acedido em 02/02/2017).
- NICOLA, SUSANA, Eduarda Pinto Ferreira e J Pinto Ferreira (2012). «A novel framework for modeling value for the customer, an essay on negotiation». Em: *International journal of information technology decision making* 11, pp. 661–703. issn: 0219-6220. doi: 10.1142/s0219622012500162.
- Woodall, Tony (2003). «Conceptualising 'Value for the Customer': An Attributional, Structural and Dispositional Analysis». Em: *Academy of Marketing Science Review* 12.5, pp. 1–42. issn: 14705931.
- Lindgreen, Adam e Finn Wynstra (2005). «Value in business markets: What do we know? Where are we going?» Em: *Industrial Marketing Management* 34.7 SPEC. ISS. Pp. 732–748. issn: 00198501. doi: 10.1016/j.indmarman.2005.01.001.
- Nylund, a (1999). «Tracing the BI family tree». Em: *Knowledge Management* 7.July, p. 2.
- Gibson, Marcus, David Arnott e Ilona Jagielska (2004). «Evaluating the Intangible Benefits of Business Intelligence: Review & Research Agenda». Em: *IFIP International Conference on Decision Support Systems (DSS2004): Decision Support in an Uncertain and Complex World*, pp. 295–305.
- Turban, Efraim, Ramesh Sharda e Dursun Delen (2010). *Decision support and business intelligence systems*. Prentice Hall. isbn: 9780136107293.
- Chee, Timothy et al. (2009b). «Business Intelligence Systems: State-of-the-art review and contemporary applications». Em: *Symposium on Progress in Information and Communication Technology*, pp. 96–101. url: http://spict.utar.edu.my/SPICT-09CD/contents/pdf/SPICT09_A-5_1.pdf.
- Chaudhuri, Surajit, Umeshwar Dayal e Vivek Narasayya (2011). «An overview of business intelligence technology». Em: *Communications of the ACM* 54.8, p. 88. issn: 00010782. doi: 10.1145/1978542.1978562.
- Kimball, Ralph e Margy Ross (2015). *The Data Warehouse Toolkit*. Ed. por Robert Elliott. Third Edit. Vol. 1. Wiley, p. 601. isbn: 9788578110796. doi: 10.1017/CB09781107415324.004. arXiv: arXiv:1011.1669v3.
- Inmon, William Harvey (2002). *Building the data warehouse*. 3, p. 427. isbn: 0471081302. url: http://cs5937.userapi.com/u11728334/docs/1ef816d44d3f/W_H_Inmon_Building_the_Data_Warehouse_3rd_E.pdf.
- Tremblay, Monica Chiarini, Alan R Hevner e Donald J Berndt (2012). «Design of an Information Volatility Measure for Health Care Decision Making». Em: *Decis. Support Syst.* 52.2, pp. 331–341. issn: 0167-9236. doi: 10.1016/j.dss.2011.08.009. url: <http://dx.doi.org/10.1016/j.dss.2011.08.009>.
- IBM (2013). «Data-driven healthcare organizations use big data analytics for big gains». Em: pp. 1–8. url: http://www03.ibm.com/industries/ca/en/healthcare/documents/Data_driven_healthcare_organizations_use_big_data_analytics_for_big_gains.pdf.
- Hanson, Ralph M (2011). «Good health information—an asset not a burden!» Em: *Australian health review : a publication of the Australian Hospital Association* 35.1, pp. 9–13. issn: 0156-5788. doi: 10.1071/AH09865. url: http://www.publish.csiro.au/view/journals/dsp_journal_fulltext.cfm?nid=270&f=AH09865.
- Coddington, Dean C e Keith D Moore (2012). «Integrating physician perspectives into business intelligence.» Em: *Healthcare Financial Management: Journal Of The Healthcare*

- Financial Management Association* 66.6, p. 158. issn: 0735-0732. url: <http://search.ebscohost.com/login.aspx?direct=true&db=mnh&AN=22734332&lang=pt-br&site=ehost-live&scope=site>.
- Data, Healthcare et al. (2014). «It All Starts with a Data Warehouse». Em: p. 74.
- Chee, Timothy et al. (2009a). «Business Intelligence Systems : State-of-the-Art Review and Contemporary Applications». Em: pp. 96–101.
- Evans, R Scott, James F Lloyd e Lee A Pierce (2012). «Clinical Use of an Enterprise Data Warehouse». Em: *AMIA Annual Symposium Proceedings* 2012, pp. 189–198. issn: 1942-597X. url: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3540441/>.
- Serviços Partilhados do Ministério da Saúde (2017). *SIM@SNS - SPMS*. url: <http://spms.min-saude.pt/product/simsns/> (acedido em 25/02/2017).
- Barrento, Manuel Pedro et al. (2010). «Business intelligence applied to Homogeneous Diagnostic Groups». Em: ... (*CISTI*), *2010 5th* ... Pp. 1–5. url: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5556653.
- Kao, Hao-Yun et al. (2016). «Design and evaluation of hospital-based business intelligence system (HBIS): A foundation for design science research methodology». Em: *Computers in Human Behavior* 62.April, pp. 495–505. issn: 07475632. doi: 10.1016/j.chb.2016.04.021. url: <http://linkinghub.elsevier.com/retrieve/pii/S0747563216302965>.
- Mann, Susan et al. (2006). «Assessing quality obstetrical care: development of standardized measures.» Em: *Joint Commission journal on quality and patient safety / Joint Commission Resources* 32.9, pp. 497–505. issn: 15537250. doi: 10.1016/S1553-7250(06)32065-X.
- WHO, UNFPA e UNICEF (2009). «AMDD: Monitoring emergency obstetric care: a handbook». Em: *Geneva: WHO* 152.4, p. 430. issn: 0144-3615. doi: 10.3109/01443611003791730. arXiv: arXiv:1011.1669v3. url: <http://informahealthcare.com/doi/abs/10.3109/01443611003791730>.
- Long, Qian et al. (2015). «Caesarean section rates in Mozambique.» Em: *BMC pregnancy and childbirth* 15, p. 253. issn: 1471-2393. doi: 10.1186/s12884-015-0686-x. url: https://apps.webofknowledge.com/full_record.do?product=UA&search_mode=CombineSearches&qid=9&SID=S1oTRdozb2JfTMMJ2DR&page=1&doc=1&cacheurlFromRightClick=no.
- Gartner, Inc (2017). *About Gartner*. url: <http://www.gartner.com/technology/about.jsp> (acedido em 14/02/2017).
- Parenteau, Josh et al. (2016). «Magic Quadrant for Business Intelligence and Analytics Platforms». Em: *Gartner Research* 1.February, pp. 1–78. issn: 1098-6596. doi: 10.1017/CB09781107415324.004. arXiv: arXiv:1011.1669v3. url: <https://www.gartner.com/doc/reprints?id=1-2XXET8P&ct=160204>.
- Predictive Analytics Today (2017). *40 Open Source and Free Business Intelligence Solutions - Predictive Analytics Today*. url: <http://www.predictiveanalyticstoday.com/open-source-free-business-intelligence-solutions/> (acedido em 16/02/2017).
- Microsoft (2017d). *Preços para o Power BI | Microsoft Power BI*. url: <https://powerbi.microsoft.com/pt-br/pricing/> (acedido em 27/05/2017).
- The Eclipse Foundation (2014). *About*. url: <http://www.eclipse.org/birt/about/> (acedido em 16/02/2017).
- SpagoBI (2017). *business intelligence | spagobi*. url: <https://www.spagobi.org/homepage/product/business-intelligence/> (acedido em 16/02/2017).
- Vaidya, Omkarprasad S. e Sushil Kumar (2006). «Analytic hierarchy process: An overview of applications». Em: *European Journal of Operational Research* 169.1, pp. 1–29. issn: 03772217. doi: 10.1016/j.ejor.2004.04.028.

- Saaty, Thomas L. (1990). «How to make a decision: The analytic hierarchy process». Em: *European Journal of Operational Research* 48.1, pp. 9–26. issn: 03772217. doi: 10.1016/0377-2217(90)90057-I. url: <http://linkinghub.elsevier.com/retrieve/pii/037722179090057I>.
- Microsoft (2017a). *Create and manage relationships in Power BI Desktop* | Microsoft Power BI. url: <https://powerbi.microsoft.com/en-us/documentation/powerbi-desktop-create-and-manage-relationships/> (acedido em 11/09/2017).
- (2017b). *How should I collaborate and share dashboards and reports?* | Microsoft Power BI. url: <https://powerbi.microsoft.com/en-us/documentation/powerbi-service-how-should-i-share-my-dashboard/> (acedido em 11/09/2017).
- (2017c). *Power BI Security Whitepaper - May 2017*.
- Liu, Shiliang et al. (2007). «Maternal mortality and severe morbidity associated with low-risk planned cesarean delivery versus planned vaginal delivery at term». Em: *Cmaj* 176.4, pp. 455–460. issn: 08203946. doi: 10.1503/cmaj.060870.
- campos, Diogo Ayres-de (2014). «A evolução da taxa de cesarianas». Em:
- Mylonas, Ioannis e Klaus Frieze (2015b). «Indikationen, Vorz??ge und Risiken einer elektiven Kaiserschnittoperation». Em: *Deutsches Arzteblatt International* 112.29-30, pp. 489–495. issn: 18660452. doi: 10.3238/arztebl.2015.0489.
- Khalil, A. et al. (2013). «Maternal age and adverse pregnancy outcome: A cohort study». Em: *Ultrasound in Obstetrics and Gynecology* 42.6, pp. 634–643. issn: 09607692. doi: 10.1002/uog.12494.
- Europeia, União (2016). «Regulamento (UE) 2016/679 d - Jornal Oficial L 119». Em:

Apêndice A

Método AHP para escolha da ferramenta

Para a escolha da ferramenta que mais se aproxima dos requisitos do cliente foi utilizado o método AHP onde as ferramentas foram avaliadas para cada critério definido. As ferramentas foram comparadas aos pares em cada critério, utilizando-se uma escala com um valor máximo de 9 (escala de Saaty).

Cada tabela apresenta a avaliação efetuada aos respectivos critérios, e a respetiva prioridade relativa (calculada pela média aritmética dos valores de cada linha da matriz normalizada).

A tabela A.1 apresenta a avaliação efetuada às alternativas para a escalabilidade. A escalabilidade representa a capacidade da ferramenta se adaptar a mais funcionalidades, por exemplo.

Tabela A.1: Avaliação das ferramentas para o critério escalabilidade.

	Pentaho	Jaspersoft	Birt	SpagoBI	Power BI	Prioridade relativa
Pentaho	1	2	1	1	$\frac{1}{2}$	0,1818
Jaspersoft	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	0,0909
Birt	1	2	1	1	$\frac{1}{2}$	0,1818
SpagoBI	1	2	1	1	$\frac{1}{2}$	0,1818
Power BI	2	4	2	2	1	0,3636

De todas as ferramentas a que possui um valor mais baixo é o Jaspersoft, uma vez que o *designer de reports* desta ferramenta não suporta a junção de várias fontes. Desta forma, a adaptação da ferramenta a novos sistemas e novas funcionalidades que necessitassem de diferentes fontes poderia estar comprometido.

De todas, a que se destaca é o Power BI não só pela integração a variadas fontes como pela facilidade de utilização, permitindo que facilmente se possa crescer na área de negócio.

A tabela A.2 apresenta os valores respetivos à comparação das alternativas para o critério da configuração. A configuração representa a capacidade de definir um conjunto de parâmetros para o funcionamento da ferramenta.

Tabela A.2: Avaliação das ferramentas para o critério configuração.

	Pentaho	Jaspersoft	Birt	SpagoBI	Power BI	Prioridade relativa
Pentaho	1	2	1	1	$\frac{1}{3}$	0,1559
Jaspersoft	$\frac{1}{2}$	1	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{5}$	0,0809
Birt	1	2	1	1	$\frac{1}{3}$	0,1559
SpagoBI	1	2	1	1	$\frac{1}{3}$	0,1559
Power BI	3	5	3	3	1	0,4514

De uma forma geral todas apresentam uma boa capacidade de configuração quer para acesso a novos dados quer para exploração. No entanto, o Power BI destaca-se por possuir linguagem natural, possuindo uma interface mais acessível, apresentando também uma maior capacidade para configuração de novos campos de análise, facilitando assim o utilizador na exploração de novos dados.

A tabela A.3 apresenta avaliação para o critério performance, desempenho da ferramenta.

Tabela A.3: Avaliação das ferramentas para o critério performance.

	Pentaho	Jaspersoft	Birt	SpagoBI	Power BI	Prioridade relativa
Pentaho	1	$\frac{1}{3}$	1	1	$\frac{1}{3}$	0,1111
Jaspersoft	3	1	3	3	1	0,3333
Birt	1	$\frac{1}{3}$	1	1	$\frac{1}{3}$	0,1111
SpagoBI	1	$\frac{1}{3}$	1	1	$\frac{1}{3}$	0,1111
Power BI	3	1	3	3	1	0,3333

De todas as alternativas as que se destacam são o Jaspersoft e o Power BI, uma vez que possuem elevadas capacidades de processamento em memória, ajudando assim a que estas possuam melhores resultados a nível de performance.

A tabela A.4 apresenta a avaliação das alternativas para o critério integração.

Tabela A.4: Avaliação das ferramentas para o critério integração.

	Pentaho	Jaspersoft	Birt	SpagoBI	Power BI	Prioridade relativa
Pentaho	1	1	1	1	$\frac{1}{2}$	0,1667
Jaspersoft	1	1	1	1	$\frac{1}{2}$	0,1667
Birt	1	1	1	1	$\frac{1}{2}$	0,1667
SpagoBI	1	1	1	1	$\frac{1}{2}$	0,1667
Power BI	2	2	2	2	1	0,3333

O Power BI destaca-se ligeiramente, porque para além das variadas ligações a diversas fontes de dados que possui, tal como as outras alternativas, possui também conexão com o facebook e google analytics.

A tabela A.5 apresenta as várias prioridades relativas que foram utilizadas para o cálculo das prioridades compostas. As prioridades compostas são obtidas pelo somatório da multiplicação da avaliação efetuada de cada alternativa em cada critério pelo peso do critério.

Tabela A.5: Apresentação das prioridades relativas e da prioridade composta.

	Escalabilidade	Configuração	Performance	Integração	Prioridade composta
Pentaho	0,1818	0,1559	0,1111	0,1667	0,1328
Jaspersoft	0,0909	0,0809	0,3333	0,1667	0,2532
Birt	0,1818	0,1559	0,1111	0,1667	0,1328
SpagoBI	0,1818	0,1559	0,1111	0,1667	0,1328
Power BI	0,3636	0,4514	0,3333	0,3333	0,3482

Para verificação de qual a melhor alternativa, a que melhor responde às necessidades do sistema, são avaliadas as prioridades compostas. A alternativa que possua a prioridade mais alta é a mais indicada para o problema em questão. A melhor alternativa para o problema apresentado é o Power BI, uma vez que é esta ferramenta que possui um valor mais alto.

Apêndice B

Inquérito apresentado aos médicos



Comentários à nova solução apresentada para o serviço de obstetrícia.

Foi desenvolvido um sistema de business intelligence aplicado aos dados obstétricos para auxiliar os profissionais desse serviço. O sistema disponibiliza os dados de forma automática e interativa. Como forma de avaliar toda a solução desenvolvida são apresentados vários pontos para que possa avaliar numa escala de 1 a 5. Onde 1 representa "Discordo completamente", 2 representa "Discordo", 3 representa "Não concordo nem discordo/não tenho opinião", 4 representa "Concordo" e 5 representa "Concordo plenamente".

	1	2	3	4	5
<i>A utilização da aplicação é intuitiva e simples.</i>					
<i>Os filtros apresentados são pertinentes para a realidade do hospital.</i>					
<i>A aplicação responde rapidamente quando aplicados filtros.</i>					
<i>A disponibilização da informação neste formato vai permitir analisar os indicadores de forma mais eficiente em particular o das taxas de cesarianas.</i>					
<i>Os indicadores apresentados são adequados ao trabalho desenvolvido neste hospital.</i>					
<i>Prefiro este método de apresentação de indicadores ao invés de relatórios estáticos.</i>					
<i>Comparativamente com a solução atual, existem melhorias com a utilização desta nova solução.</i>					
<i>Gostaria de ter esta ferramenta de exploração de dados no seu Serviço de Obstetrícia.</i>					

Acrescentaria mais algum indicador à solução? Se sim, quais?

Apêndice C

Inquérito apresentado aos colaboradores da VirtualCare



Comentários técnicos à solução desenvolvida para o serviço de obstetrícia.

Foi desenvolvido um sistema de business intelligence aplicado aos dados obstétricos para auxiliar os profissionais desse serviço. O sistema disponibiliza os dados de forma automática e interativa. Como forma de avaliar toda a solução desenvolvida são apresentados vários pontos para que possa avaliar numa escala de 1 a 5. Onde 1 representa "Discordo completamente", 2 representa "Discordo", 3 representa "Não concordo nem discordo/não tenho opinião", 4 representa "Concordo" e 5 representa "Concordo plenamente".

	1	2	3	4	5
<i>A arquitetura da solução é adequada às necessidades do negócio.</i>					
<i>A arquitetura desenvolvida permite crescer para novos indicadores</i>					
<i>A aplicação é facilmente integrada com outros sistemas.</i>					
<i>Houve melhorias em relação à solução atual.</i>					